

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 10:37:17 ; Search time 116 Seconds  
(without alignments)  
1693.744 Million cell updates/sec

Title: US-09-936-702-3  
Perfect score: 2668  
Sequence: 1 MNRGVPFRHLVLQLALP.....GTRLEIKLVPRGSGHHHHH 508

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2651.5	99.4	507	3 AAB00158	AAB00158 sCD4-SCFv
2	1064	39.9	295	2 AAR07605	AAR07605 Plasmid T
3	1064	39.9	318	2 AAR07605	AAR07605 Plasmid T
4	1064	39.9	394	2 AAY39825	AAY39825 Soluble h
5	1064	39.9	394	3 AAY88328	AAY88328 T4 glycop
6	1064	39.9	416	3 AAB19509	Aab19509 CD4-IgM f
7	1064	39.9	436	3 AAY51080	AAY51080 Human fus
8	1064	39.9	458	3 AAY88329	Aay88329 T4 glycop
9	1064	39.9	458	4 AAB81502	Aab81502 Human CD4
10	1064	39.9	458	4 AAG79087	Aag79087 Amino aci
11	1064	39.9	458	7 ADD25609	Add25609 Binding d
12	1064	39.9	458	7 ADE57489	Ade57489 Human Pro
13	1064	39.9	458	8 ADP12495	Adp12495 Protein e
14	1064	39.9	473	7 ADA44807	Ada44807 CD4/TCR C
15	1064	39.9	474	3 AAY59170	Aay59170 CD4-Ig fu
16	1064	39.9	481	3 AAB19510	Aab19510 CD4-IgM f
17	1064	39.9	481	3 AAY51081	Aay51081 Human fus
18	1064	39.9	481	3 AAY59171	Aay59171 CD4-Ig fu
19	1064	39.9	519	2 AAR20152	AAR20152 Human CD4
20	1064	39.9	616	3 AAY51082	Aay51082 Human fus
21	1064	39.9	616	3 AAY59172	Aay59172 CD4-Ig fu
22	1064	39.9	631	1 AAP93009	Aap93009 Genetic c
23	1064	39.9	631	3 AAB19508	Aab19508 CD4-IgG1
24	1064	39.9	631	3 AAY51079	Aay51079 Human fus
25	1064	39.9	631	3 AAY59169	Aay59169 CD4-Ig fu

26	1064	39.9	729	1 AAP93008	Aap93008 Genetic c
27	1064	39.9	729	3 AAB19507	Aab19507 CD4-IgG1
28	1064	39.9	729	3 AAY51078	Aay51078 Human fus
29	1064	39.9	729	3 AAY59168	Aay59168 CD4-Ig fu
30	1062	39.8	458	7 ADE65841	Ade65841 Human CD4
31	1059	39.7	398	2 AAR78673	Aar78673 CD4 doma
32	1059	39.7	398	2 AAR89450	Aar89450 CD4 DI-D4
33	1059	39.7	400	2 AAR06374	Aar06374 Truncated
34	1059	39.7	402	1 AAP91922	Aap91922 Sequence
35	1059	39.7	402	1 AAP94757	Aap94757 Sequence
36	1059	39.7	458	1 AAP81990	Aap81990 Clone pT4
37	1059	39.7	458	1 AAP91369	Aap91369 T4 protei
38	1059	39.7	458	2 AAR13491	Aar13491 Human CD4
39	1059	39.7	458	2 AAY39826	Aay39826 Soluble h
40	1059	39.7	458	3 AAB07769	Aab07769 DNA encod
41	1059	39.7	462	2 AAR27277	Aar27277 CD4:eta p
42	1059	39.7	462	2 AAR78677	Aar78677 T-cell re
43	1059	39.7	462	2 AAR89457	Aar89457 CD4:gamma
44	1059	39.7	462	2 AAW02214	Aaw02214 CD4:Fc re
45	1059	39.7	462	2 AAW83142	Aaw83142 Chimeric

ALIGNMENTS

RESULT 1	
ID	AAB00158 standard; protein; 507 AA.
XX	
AC	AAB00158;
XX	
DT	08-FEB-2001 (first entry)
XX	
DE	sCD4-SCFv(17b) HIV single chain antibody fusion protein.
XX	
KW	Fusion protein; HIV; human immunodeficiency virus; antibody; Fv; AIDS; -
KW	acquired immune deficiency syndrome; neutralisation; infection;
KW	gene therapy; CD4; gp120; glycoprotein; resistance; vaccination;
KW	binding domain; single chain antibody; chimera; chimeric protein.
XX	
OS	Human immunodeficiency virus.
OS	Synthetic.
XX	
PN	MO200055207-A1.
XX	
PD	21-SEP-2000.
XX	
PF	16-MAR-2000; 2000MO-US006946.
XX	
PR	16-MAR-1999; 99US-0124681P.
XX	
PA	(USSH ) US NAT INST OF HEALTH.
XX	
PI	Berger EA, Del Caetlillo CM;
XX	
DR	WPI: 2000-638183/61.
DR	N-PSDB; AAA54045.
XX	
PT	Novel neutralizing bispecific fusion proteins effective in viral such as
PT	HIV neutralization, comprises two different binding domains, inducing-
PT	binding domain and induced-binding domain functionally linked by linker.
XX	
PS	Claim 39; Page 46-47; 55pp; English.
XX	
CC	sCD4-SCFv(17b) is a neutralising bispecific fusion protein capable of
CC	binding to two sites of its target protein. The protein comprises a first
CC	binding domain capable of binding to an inducing site on the target
CC	protein, a second binding domain capable of forming neutralising complex
CC	with an induced epitope of the target protein and a linker connecting the
CC	binding domains. sCD4-SCFv(17b) comprises a soluble CD4 fragment
CC	(containing domains D1 and D2) fused to a single chain Fv portion of
CC	antibody 17b via a linker. sCD4-SCFv(17b), its variant, analogue or
CC	mimetic is used for inactivating gp120 protein of HIV, and for

CC neutralising HIV. It is also used for blocking and preventing the binding  
 CC of the viral or recombinant gp120 protein to soluble CD4 or lymphocyte  
 CC CD4 and for inhibiting HIV replication. The chimeric proteins is  
 CC therefore useful for treating HIV infection and also AIDS. It is are  
 CC particularly useful in the prevention of infection during or immediately  
 CC after HIV exposure (e.g., mother/infant transmission, post-exposure  
 CC prophylaxis, and as a topical inhibitor) and for providing long term  
 CC resistance to HIV infections and AIDS. Gene therapy is used to secrete  
 CC the bispecific protein at mucosal surfaces, such as the vaginal, rectal  
 CC or oral mucosa. The fusion proteins is highly potent, broadly cross-  
 CC reactive with neutralising antibody with high in vivo activity and no Fc-  
 CC mediated undesirable targeting properties. When the fusion protein is  
 CC substantially derived from human proteins, it has minimal immunogenicity  
 CC and toxicity in humans which is of great value in prevention of infection  
 CC during or immediately after HIV exposure

XX  
 XX Sequence 507 AA;

Query Match 99.4%; Score 2651.5; DB 3; Length 507;  
 Best Local Similarity 99.8%; Pred. No. 9.4e-128;  
 Matches 507; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MNRGVPFRHLLVQLALLPAATQGGKVVLGKKDVELTCTASOKKSIQFHKNSNOIK 60  
 DB 1 MNRGVPFRHLLVQLALLPAATQGGKVVLGKKDVELTCTASOKKSIQFHKNSNOIK 60  
 QY 61 ILNGQSFLLTKGSPKLNDRADSRSLMDQGNFLLIKNLKI ESDTYICEVEQKEEVOL 120  
 DB 61 ILNGQSFLLTKGSPKLNDRADSRSLMDQGNFLLIKNLKI ESDTYICEVEQKEEVOL 120  
 QY 121 LVFGLTANSDTHLLQCGSLTLTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 DB 121 LVFGLTANSDTHLLQCGSLTLTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 QY 181 TWCTCVLQNKKEFKIDIVLAFOKASGGGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 240  
 DB 181 TWCTCVLQNKKEFKIDIVLAFOKASGGGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 240  
 QY 241 GSGSQVQLLESQAEVKKPSSVKVSCKASGDTFIRYFTWVROAPGQLEMMGRITTLDV 300  
 DB 241 GSGSQVQLLESQAEVKKPSSVKVSCKASGDTFIRYFTWVROAPGQLEMMGRITTLDV 300  
 QY 301 AHYAPHLQGRVTTTADKSTSTVLELNLRASDDTAVYFCAGVYEGEADDEGEYDNNGLKH 360  
 DB 301 AHYAPHLQGRVTTTADKSTSTVLELNLRASDDTAVYFCAGVYEGEADDEGEYDNNGLKH 360  
 QY 361 WGGGLTVTVSGGGSGSGSGSGSGSGSBLTQSPATLSVSGERATLSCPASSESVSDLA 420  
 DB 361 WGGGLTVTVSGGGSGSGSGSGSGSBLTQSPATLSVSGERATLSCPASSESVSDLA 420  
 QY 421 WYQOKPGQAPRLILYGASTRATGVPARFSGSGSAGEFTLTITISLQSEDFAVYYCOQYNNW 480  
 DB 421 WYQOKPGQAPRLILYGASTRATGVPARFSGSGSAGEFTLTITISLQSEDFAVYYCOQYNNW 480  
 QY 481 PPRYTFGGQTRLEIKLVPRGSGHHHHH 508  
 DB 481 PPRYTFGGQ-TRLEIKLVPRGSGHHHHH 507

RESULT 2  
 AAR07605  
 ID AAR07605 standard; protein; 295 AA.  
 XX  
 XX AAR07605;  
 AC  
 XX 25-MAR-2003 (revised)  
 DT 09-JAN-2003 (revised)  
 DT 06-FEB-1991 (first entry)  
 XX  
 XX Plasmid T4/LFA-3/AD gene product with C-terminal encoding  
 DE phosphatidylinositol (PI) linkage signaling sequence.  
 XX  
 XX Plasma membrane binding affinity; micelle.

XX Unidentified.  
 OS  
 XX Key Location/Qualifiers  
 FH Protein 1..266  
 FT Peptide /label= T4/LFA-3/AD gene product  
 FT Peptide 267..295  
 FT /label= PI signal peptide

PN W09012099-A.  
 XX  
 PD 18-OCT-1990.  
 XX  
 PF 10-APR-1989; 89US-00335688.  
 XX  
 PR 10-APR-1989; 89US-00335688.  
 XX  
 PA (BIOU) BIOGEN INC.  
 XX  
 PI Wallner BP;  
 XX  
 DR WPI: 1990-334849/44.  
 XX  
 DR N-PSDB; AAQ06404.  
 XX  
 PT Phosphatidy1-inositol linkage signalling DNA sequence - derived from  
 PT lymphocyte function-associated antigen 3, used for prodn. of chimeric  
 PT proteins.  
 XX  
 PS Disclosure; Fig 4; 53pp; English.  
 XX  
 CC The signal sequence is attached downstream of the plasmid sequence  
 CC encoding a secreted protein, which will then produce proteins covalently  
 CC anchored to the cell surface in which they were produced. This can give  
 CC rise to plasma membrane binding, enhanced purifiability, micelle  
 CC formation etc. especially useful in the production of chimeric targeted  
 CC drugs; to produce micellar or liposomal delivery systems or in enhanced  
 CC purification and screening of cells, proteins or DNA libraries. (updated  
 CC on 09-JAN-2003 to add missing OS field.) (updated on 25-MAR-2003 to  
 CC correct PA field.)  
 XX  
 XX Sequence 295 AA;

Query Match 39.9%; Score 1064; DB 2; Length 295;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALLPAATQGGKVVLGKKDVELTCTASOKKSIQFHKNSNOIK 60  
 DB 1 MNRGVPFRHLLVQLALLPAATQGGKVVLGKKDVELTCTASOKKSIQFHKNSNOIK 60  
 QY 61 ILNGQSFLLTKGSPKLNDRADSRSLMDQGNFLLIKNLKI ESDTYICEVEQKEEVOL 120  
 DB 61 ILNGQSFLLTKGSPKLNDRADSRSLMDQGNFLLIKNLKI ESDTYICEVEQKEEVOL 120  
 QY 121 LVFGLTANSDTHLLQCGSLTLTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 DB 121 LVFGLTANSDTHLLQCGSLTLTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 QY 181 TWCTCVLQNKKEFKIDIVLAFOKAS 208  
 DB 181 TWCTCVLQNKKEFKIDIVLAFOKAS 208

RESULT 3  
 AAR07606  
 ID AAR07606 standard; protein; 318 AA.  
 XX  
 XX AAR07606;  
 AC  
 XX 25-MAR-2003 (revised)  
 DT 09-JAN-2003 (revised)  
 DT 06-FEB-1991 (first entry)  
 XX

DE Plasmid T4/LFA-3/2 gene product with C-terminal encoding  
 DE phosphatidylinositol (PI) linkage signaling sequence.  
 XX  
 KM Plasma membrane binding affinity; micelle.  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 1..267  
 FT /label= LFA-3 gene product  
 FT Peptide 267..318  
 FT /label= PI signal peptide  
 XX  
 PN MO9012099-A.  
 XX  
 PD 18-OCT-1990.  
 XX  
 PF 10-APR-1989; 89US-00335688.  
 XX  
 PR 10-APR-1989; 89US-00335688.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 XX  
 PI Wallner BP;  
 XX  
 PI WPI; 1990-334849/44.  
 DR N-PSDB; AAQ06405.  
 XX  
 PT Phosphatidylinositol linkage signalling DNA sequence - derived from  
 PT lymphocyte function-associated antigen 3, used for prodn. of chimeric  
 PT proteins.  
 XX  
 PS Disclosure; Fig 5; 53pp; English.  
 XX  
 CC The signal sequence is attached downstream of the Plasmid sequence  
 CC encoding a secreted protein, which will then produce proteins covalently  
 CC anchored to the cell surface, in which they were produced. This can give  
 CC rise to plasma membrane binding, enhanced purifiability, micelle  
 CC formation etc. especially useful in the production of chimeric targeted  
 CC drugs; to produce micellar or liposomal delivery systems or in enhanced  
 CC purification and screening of cells, proteins or DNA libraries. (Updated  
 CC on 09-JAN-2003 to add missing OS field.) (Updated on 25-MAR-2003 to  
 CC correct PA field.)  
 CC  
 XX  
 SQ Sequence 318 AA;  
 Query Match 39.9%; Score 1064; DB 2; Length 318;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MNRGVPFRHLLVLTQALLPATQGGKVVLTGKGDVETLCTASOKKSIOFHKNSNOIK 60  
 DB 1 MNRGVPFRHLLVLTQALLPATQGGKVVLTGKGDVETLCTASOKKSIOFHKNSNOIK 60  
 QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120  
 DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120  
 QY 121 LVFGLTANSDTHLLQGSQSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180  
 DB 121 LVFGLTANSDTHLLQGSQSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180  
 QY 181 TWTCVTIVONOKKVEFKIDIVLAFQKAS 208  
 DB 181 TWTCVTIVONOKKVEFKIDIVLAFQKAS 208  
 RESULT 4  
 ID AAY39825 standard; protein; 394 AA.  
 XX  
 AC AAY39825;  
 XX

DT 03-DEC-1999 (first entry)  
 XX  
 DE Soluble human T4 protein.  
 XX  
 KM Soluble T4 protein; sT4; human; HIV; binding inhibitor; T4+ cell; AIDS;  
 KM vaccine; immunisation; therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN USS958678-A.  
 XX  
 PD 28-SEP-1999.  
 XX  
 PF 12-DEC-1994; 94US-00354452.  
 XX  
 PR 21-AUG-1986; 86US-00898587.  
 PR 11-JUN-1991; 91US-00713564.  
 PR 06-JUL-1992; 92US-00909021.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PI McDougal JS, Weiss R, Axel R, Littman DR, Maddon PJ, Chess L;  
 XX  
 PI WPI; 1999-561025/47.  
 DR N-PSDB; AAZ20694.  
 XX  
 PT Human T4 protein inhibits HIV binding to T4 cells, useful for treating  
 PT AIDS.  
 XX  
 PS Disclosure; Col 13-16; 56pp; English.  
 XX  
 CC This sequence represents the soluble human T4 protein of the invention.  
 CC The soluble human T4 protein blocks the binding of HIV to T4+ cells and  
 CC is therefore useful for the treatment of AIDS. Monoclonal antibodies  
 CC against the T4 protein may be used as vaccines for immunising subjects  
 CC against AIDS  
 CC  
 XX  
 SQ Sequence 394 AA;  
 Query Match 39.9%; Score 1064; DB 2; Length 394;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MNRGVPFRHLLVLTQALLPATQGGKVVLTGKGDVETLCTASOKKSIOFHKNSNOIK 60  
 DB 1 MNRGVPFRHLLVLTQALLPATQGGKVVLTGKGDVETLCTASOKKSIOFHKNSNOIK 60  
 QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120  
 DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120  
 QY 121 LVFGLTANSDTHLLQGSQSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180  
 DB 121 LVFGLTANSDTHLLQGSQSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180  
 QY 181 TWTCVTIVONOKKVEFKIDIVLAFQKAS 208  
 DB 181 TWTCVTIVONOKKVEFKIDIVLAFQKAS 208  
 RESULT 5  
 ID AAY88328 standard; protein; 394 AA.  
 XX  
 AC AAY88328;  
 XX  
 DT 14-JUL-2000 (first entry)  
 XX  
 DE T4 glycoprotein amino acid sequence.  
 XX  
 KM sT4; glycoprotein; human immunodeficiency virus; HIV; block binding;  
 KM AIDS; treatment; inhibit; cell to cell spread; infection; fusion.  
 XX

```

OS Mammalia.
XX US5126433-A.
XX 30-JUN-1992.
XX
XX 23-OCT-1987; 87US-00114244.
XX
XX 21-AUG-1986; 86US-00898587.
XX
XX (UYCO ) UNIV COLUMBIA NEW YORK.
XX
XX Maddon PJ, Chess L, Axel R, Weiss R, Littman DR, McDougal JS;
XX WPI; 2000-348913/30.
XX N-PSDB; AAA10906.
XX
XX Soluble T4 glycoprotein useful for prevention and treatment of acquired
XX immunodeficiency syndrome and for screening inhibitors of human
XX immunodeficiency viral binding.
XX
XX Disclosure; Col 11-16; 64pp; English.
XX
XX This sequence represents the full length amino acid sequence of
XX glycosylated sT4 glycoprotein. Human immunodeficiency virus (HIV) uses
XX sT4 as a target receptor on T cells. The invention relates to
XX glycosylated sT4 which functions by blocking the binding of HIV to T4
XX target cells, and can be used for the prophylaxis and treatment of AIDS
XX patients. Administration of sT4 effectively inhibits the cell to cell
XX spreading of HIV infection and also the fusion of HIV-infected T4 cells
XX and non-infected T4 cells. The administration of T4 alleviates several
XX symptoms associated with AIDS, and prevents the occurrence of new
XX pathological changes. The sT4 glycoprotein is useful for the prophylaxis
XX and treatment of patients with AIDS. It is also useful as a reagent to
XX identify natural, synthetic or recombinant molecules which act as
XX therapeutic agents or inhibitors of T4+ cell interactions and in
XX diagnostic assays for detection T4 proteins or molecules
XX
XX Sequence 394 AA:
SQ
Query Match 39.9%; Score 1064; DB 3; Length 394;
Best Local Similarity 100.0%; Pred. No. 8.3e-47;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVFRRLLVLTQALPAATQGGKVVLGKGDVVELTCTASQKKSIQFHKNSNOIK 60
DB 1 MNRGVFRRLLVLTQALPAATQGGKVVLGKGDVVELTCTASQKKSIQFHKNSNOIK 60
QY 61 ILGNQGSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEVOL 120
DB 61 ILGNQGSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEVOL 120
QY 121 LVFGLTANSPTHTLLOGQSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSPTHTLLOGQSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCTVLQNGKVEFKIDIVLAFQKAS 208
DB 181 TWTCTVLQNGKVEFKIDIVLAFQKAS 208

RESULT 6
AAB19509
ID AAB19509 standard; protein; 416 AA.
XX
XX AAB19509;
XX
XX 09-JAN-2001 (first entry)
XX
XX CD4-IgM fusion protein CH4mmu.
XX
XX CD4; IgM, human; CD4mmu; fusion protein; immunoglobulin; HIV; SIV; gp120;
XX therapy; diagnosis.

```

```

XX OS Homo sapiens.
XX Key Location/Qualifiers
XX FH 1..395"CD4 extracellular region"
XX FT /note="CD4 extracellular region"
XX FT Protein 400..416
XX FT /note="IgM heavy chain partial sequence"
XX
XX US6117656-A.
XX
XX PD 12-SEP-2000.
XX
XX PF 07-JUN-1995; 95US-00479353.
XX
XX 22-JAN-1988; 88US-00147351.
XX PR 23-JAN-1989; 89US-00299596.
XX PR 09-JUN-1992; 92US-00895781.
XX PR 12-APR-1993; 93US-00057952.
XX PR 04-FEB-1994; 94US-00191708.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX
XX Seed B;
XX
XX WPI; 2000-586558/55.
XX N-PSDB; AAA50662.
XX
XX CD4-immunoglobulin fusion proteins, useful for targeting gp120 of HIV or
XX SIV.
XX
XX Example 1; Col 41-50; 39pp; English.
XX
XX The present sequence is that of fusion protein CD4mmu comprising the
XX extracellular portion of CD4, which binds to HIV gp120, linked at its C-
XX terminus to the human IgM heavy chain. To obtain the fusion protein, DNA
XX encoding CD4 was linked to IgM DNA at the MscI site upstream of the CH1
XX region (see AAA50662). Fusion protein CD4mmu and a nucleic acid encoding
XX it are claimed. Also claimed are a vector comprising the nucleic acid,
XX a method of producing the fusion protein in secreted form using a
XX transformed host cell. The fusion protein may further comprise a
XX therapeutic agent, radiolabel or NMR imaging agent. The fusion protein
XX can be administered to an animal (including humans) for treatment of HIV
XX or SIV infection, and can also be used in assays for HIV or SIV, imaging
XX and tissue stains. IgM fusion proteins such as CD4mmu provide complement-
XX mediated immunity
XX
XX Sequence 416 AA:
SQ
Query Match 39.9%; Score 1064; DB 3; Length 416;
Best Local Similarity 100.0%; Pred. No. 8.7e-47;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVFRRLLVLTQALPAATQGGKVVLGKGDVVELTCTASQKKSIQFHKNSNOIK 60
DB 1 MNRGVFRRLLVLTQALPAATQGGKVVLGKGDVVELTCTASQKKSIQFHKNSNOIK 60
QY 61 ILGNQGSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEVOL 120
DB 61 ILGNQGSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEVOL 120
QY 121 LVFGLTANSPTHTLLOGQSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSPTHTLLOGQSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCTVLQNGKVEFKIDIVLAFQKAS 208
DB 181 TWTCTVLQNGKVEFKIDIVLAFQKAS 208

RESULT 7
AAY51080
ID AAY51080 standard; protein; 436 AA.

```



XX AAY51080;  
 AC  
 XX  
 DT 23-MAR-2000 (first entry)  
 XX  
 DE Human fusion protein CD4Mg.  
 XX  
 DE Fusion protein; human; CD4; IgM; immunoglobulin; gp120;  
 KM anti-human immunodeficiency virus; CD4Mg.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 PN US6004781-A.  
 XX  
 PD 21-DEC-1999.  
 XX  
 PF 04-FEB-1994; 94US-00191708.  
 XX  
 PR 22-JAN-1988; 88US-00147351.  
 PR 23-JAN-1989; 89US-00299596.  
 PR 09-JUN-1992; 92US-00896781.  
 PR 12-APR-1993; 93US-00057952.  
 XX  
 PA (GEHO ) GEN HOSPITAL CORP.  
 XX  
 PI Seed B;  
 XX  
 DR WPI; 2000-085792/07.  
 DR N-PSDB; AAZ44063.  
 XX  
 PT Fusion protein useful for the treatment of human immunodeficiency virus.  
 XX  
 PS Example 1; Col 41-50; 39pp; English.  
 CC This invention describes a novel nucleic acid (I) encoding a fusion  
 CC protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)  
 CC and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light  
 CC chain (III). The products of the invention have anti-human  
 CC immunodeficiency virus (HIV) activity and are capable of binding to  
 CC gp120. The fusion protein is useful for treating human immunodeficiency  
 CC virus (HIV) or simian immunodeficiency virus (SIV). This sequence  
 CC represents the fusion protein CD4Mg which is constructed from CD4 linked  
 CC to human Igm upstream of the CH1 region  
 CC  
 XX  
 SQ Sequence 436 AA;  
 Query Match 39.9%; Score 1064; DB 3; Length 436;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALLPAATOGKKVVGKGGDTVELTCTASOKKSIOFHMKNNSQIK 60  
 DB 1 MNRGVPFRHLLVQLALLPAATOGKKVVGKGGDTVELTCTASOKKSIOFHMKNNSQIK 60  
 QY 61 ILNGQSFLTKGPKSLNDRADSRSLMDQGNFLLIKNLKTEDSDTYICEVEDQKEEVOL 120  
 DB 61 ILNGQSFLTKGPKSLNDRADSRSLMDQGNFLLIKNLKTEDSDTYICEVEDQKEEVOL 120  
 QY 121 LVFGLTANSDTHLLQGSLLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDTHLLQGSLLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWTCIVLQNGKKEFKIDIVVLAFOKAS 208  
 DB 181 TWTCIVLQNGKKEFKIDIVVLAFOKAS 208

RESULT 8  
 AAY88329 standard; protein; 458 AA.  
 XX  
 AC AAY88329;

XX 14-JUL-2000 (first entry)  
 DT  
 XX  
 DE T4 glycoprotein amino acid sequence.  
 XX  
 DE sT4; glycoprotein; human immunodeficiency virus; HIV; block binding;  
 KM AIDS; treatment; inhibit; cell to cell spread; infection; fusion.  
 XX  
 OS Mammalia.  
 XX  
 PN US5126433-A.  
 XX  
 PD 30-JUN-1992.  
 XX  
 PF 23-OCT-1987; 87US-00114244.  
 XX  
 PR 21-AUG-1986; 86US-00898587.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PI Maddon PJ, Chess L, Axel R, Weiss R, Littman DR, McDougal JS;  
 XX  
 DR WPI; 2000-348913/30.  
 XX  
 PT Soluble T4 glycoprotein useful for prevention and treatment of acquired  
 PT immunodeficiency syndrome and for screening inhibitors of human  
 PT immunodeficiency viral binding.  
 XX  
 PS Example; Fig 6; 64pp; English.  
 CC This sequence represents the amino acid sequence of glycosylated sT4  
 CC glycoprotein. Human immunodeficiency virus (HIV) uses sT4 as a target  
 CC receptor on T cells. The invention relates to glycosylated sT4 which  
 CC functions by blocking the binding of HIV to T4 target cells, and can be  
 CC used for the prophylaxis and treatment of AIDS patients. Administration  
 CC of sT4 effectively inhibits the cell to cell spreading of HIV infection  
 CC and also the fusion of HIV-infected T4 cells and non-infected T4 cells.  
 CC The administration of T4 alleviates several symptoms associated with  
 CC AIDS, and prevents the occurrence of new pathological changes. The sT4  
 CC glycoprotein is useful for the prophylaxis and treatment of patients with  
 CC AIDS. It is also useful as a reagent to identify natural, synthetic or  
 CC recombinant molecules which act as therapeutic agents or inhibitors of  
 CC T4+ cell interactions and in diagnostic assays for detection T4 proteins  
 CC or molecules  
 CC  
 XX  
 SQ Sequence 458 AA;  
 Query Match 39.9%; Score 1064; DB 3; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALLPAATOGKKVVGKGGDTVELTCTASOKKSIOFHMKNNSQIK 60  
 DB 1 MNRGVPFRHLLVQLALLPAATOGKKVVGKGGDTVELTCTASOKKSIOFHMKNNSQIK 60  
 QY 61 ILNGQSFLTKGPKSLNDRADSRSLMDQGNFLLIKNLKTEDSDTYICEVEDQKEEVOL 120  
 DB 61 ILNGQSFLTKGPKSLNDRADSRSLMDQGNFLLIKNLKTEDSDTYICEVEDQKEEVOL 120  
 QY 121 LVFGLTANSDTHLLQGSLLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDTHLLQGSLLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWTCIVLQNGKKEFKIDIVVLAFOKAS 208  
 DB 181 TWTCIVLQNGKKEFKIDIVVLAFOKAS 208

RESULT 9  
 AAB81502 standard; protein; 458 AA.  
 XX  
 AC AAB81502;

DT	18-JUN-2001	(first entry)
XX		
DE	Human CD4 protein.	
XX		
KW	Human; CD4; CD4 fusion protein; oligomerisation; receptor-ligand interaction inhibition; surface plasmon resonance; SPR, T cell receptor binding; MHC binding; carcinoma; autoimmune disease; multiple sclerosis; human immunodeficiency virus; HIV; diabetes; rheumatoid arthritis; immune disorder.	
XX		
OS	Homo sapiens.	
XX		
FI	Key	Location/Qualifiers
FT	Peptide	1..25
FT	Protein	/label= Signal_peptide 26..458 /label= Human_CD4
XX		
PN	MO200122084-AZ.	
PD	29-MAR-2001.	
PF	18-SEP-2000; 2000WO-GB003579.	
PR	21-SEP-1999; 99GB-00022352.	
XX		
PA	(AVID -) AVIDEX LTD.	
XX		
P1	Jakobsen BK;	
XX		
DR	WPI; 2001-273470/28.	
N-PSDB; AAF82582.		
PT	Sequential screening of candidate compounds library for those which inhibit binding of low affinity receptor-ligand interaction having fast binding kinetics, using interfacial optical assay.	
XX		
PS	Disclosure; Fig 13; 91pp; English.	
XX		
CC	The present sequence is human CD4. Human CD4 extracellular domains 1 and 2 were used in the construction of CD4 oligomerisation fusion proteins. The fusion proteins contain an oligomerisation domain that enables the proteins to bind to one another to form oligomers. The oligomers may be used in an invention relating to a method for screening for compounds with the ability to inhibit a low affinity receptor-ligand interaction. The method uses an interfacial optical assay, such as surface plasmon resonance (SPR). The method is useful for screening candidate compounds for the ability to inhibit interaction between MHC/peptide complex and T cell receptor, and MHC/peptide complex and CD8 or CD4 co-receptor. The compounds identified by the above methods which interfere with T cell receptor binding to a particular HLA type molecule are useful as immune inhibitors for treating carcinomas, autoimmune diseases such as multiple sclerosis, human immunodeficiency virus (HIV) infection, rheumatoid arthritis, Hashimoto's disease, insulin dependent diabetes, Good pasture's syndrome, uveitis, psoriasis and graft rejection	
XX		
SO	Sequence 458 AA;	
Query Match	39.9%; Score 1064; DB 4; Length 458;	
Best Local Similarity	100.0%; Pred. No. 9.5e-47;	
Matches	208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MNRGVPFPHLLIVLQALLPAATQKVVLGKKSDVTEITCTASQKSIQFMKXNSNOIK 60	
DB	1 MNRGVPFPHLLIVLQALLPAATQKVVLGKKSDVTEITCTASQKSIQFMKXNSNOIK 60	
QY	ILGNQGSFLTKGSPKLTNDRAISRRLMPOGNPLIILKULKTEDSPVYCEVEDQKEEYQL 120	
DB	61 ILGNQGSFLTKGSPKLTNDRAISRRLMPOGNPLIILKULKTEDSPVYCEVEDQKEEYQL 120	
QY	121 LVFGLTANSDFTHLLQGQSLVTLTLESPGSSPSVQGRSFRGNKIQQGKTLVSQLELDQSG 180	
DB	121 LVFGLTANSDFTHLLQGQSLVTLTLESPGSSPSVQGRSFRGNKIQQGKTLVSQLELDQSG 180	

[illegible]

QY 181 TWTCVTVLQONKKVEFKIDIVLAFOKAS 208  
 |||||  
 DB 181 TWTCVTVLQONKKVEFKIDIVLAFOKAS 208

RESULT 11  
 ADD25609  
 ID ADD25609 standard; protein; 458 AA.  
 AC ADD25609;  
 XX  
 DT 15-JAN-2004 (first entry)  
 DE Binding domain-immunoglobulin fusion protein-associated protein #82.  
 XX  
 KW Binding domain; immunoglobulin; fusion protein; cytostatic;  
 KW antiarthritic; immunosuppressive; antidiabetic; antithyroid;  
 KW neuroprotective; hinge region; immunoglobulin heavy chain;  
 KW CH2 constant region; CH3 constant region; IgG1;  
 KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;  
 KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;  
 KW rheumatoid arthritis; myasthenia gravis; Grave's disease;  
 KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.  
 OS Unidentified.  
 XX  
 PN US2003118592-A1.  
 XX  
 PD 26-JUN-2003.  
 XX  
 PF 25-JUL-2002; 2002US-00207655.  
 XX  
 PR 17-JAN-2001; 2001US-0367358P.  
 PR 17-JAN-2002; 2002US-00053530.  
 PR 03-JUN-2002; 2002US-0385691P.  
 XX  
 PA (GENE-) GENE-CRAFT INC.  
 XX  
 PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;  
 DR WPI, 2003-801317/75.  
 XX  
 PT New binding domain-immunoglobulin fusion protein, useful for treating a  
 PT subject having or suspected of having a malignant condition or a B-cell  
 PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.  
 PS Disclosure; SEQ ID NO 170; 157pp; English.  
 XX  
 CC The invention relates to a binding domain-immunoglobulin fusion protein  
 CC comprising a binding domain polypeptide that is fused to an  
 CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain  
 CC CH2 constant region polypeptide that is fused to the hinge region  
 CC polypeptide, and an immunoglobulin heavy chain CH3 constant region  
 CC polypeptide that is fused to the CH2 constant region polypeptide. The  
 CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin  
 CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge  
 CC region polypeptide; derived from (a) having 3 or more cysteine residues;  
 CC where the mutated human IgG1 immunoglobulin hinge region polypeptide  
 CC contains 2 cysteine residues, where the first cysteine is not mutated; a  
 CC mutated human IgG1 immunoglobulin hinge region polypeptide; derived from  
 CC (a) having 3 or more cysteine residues, where the mutated human IgG1  
 CC immunoglobulin hinge region polypeptide contains no more than one  
 CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region  
 CC polypeptide; derived from (a) having 3 or more cysteine residues; where  
 CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains  
 CC no cysteine residues. The binding domain-immunoglobulin fusion protein is  
 CC capable of at least one immunological activity comprising antibody  
 CC dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The  
 CC binding domain polypeptide is capable of specifically binding to an  
 CC antigen. Also included are an isolated polynucleotide encoding the  
 CC binding domain-immunoglobulin fusion protein, a recombinant expression  
 CC construct comprising the polynucleotide (operably linked to a promoter),  
 CC a host cell transformed or transfected with a recombinant expression

CC construct, producing the binding domain-immunoglobulin fusion protein, a  
 CC pharmaceutical composition comprising the binding domain-immunoglobulin  
 CC fusion protein or polynucleotide and a carrier, and treating a subject  
 CC having or suspected of having a malignant condition or a B-cell disorder.  
 CC The binding domain-immunoglobulin fusion protein is useful for treating a  
 CC subject having or suspected of having a malignant condition or a B-cell  
 CC disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,  
 CC myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple  
 CC sclerosis or autoimmune disease. The present sequence is a binding domain  
 CC -immunoglobulin fusion protein-associated protein sequence. Note: The  
 CC sequence data for this patent formed part of the printed specification  
 CC and is also available in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?docID=20030118592. The authors have not  
 CC identified the sequences in the printed specification by their SEQ ID  
 CC number therefore none of the sequences can be explicitly identified.  
 XX  
 SQ Sequence 458 AA;  
 XX  
 QY 1 MNRGVPPRHLLVQLALPPATQKKVVGKKGTVELCTRASQKKSIOFHKNNSQIK 60  
 DB 1 MNRGVPPRHLLVQLALPPATQKKVVGKKGTVELCTRASQKKSIOFHKNNSQIK 60  
 QY 61 ILNGQSFLTKGSKNDRAISRSLMDQNFPLIKNLKIEPSDYICEVEPOKKEVOL 120  
 DB 61 ILNGQSFLTKGSKNDRAISRSLMDQNFPLIKNLKIEPSDYICEVEPOKKEVOL 120  
 QY 121 LVFGLTRANSDTHLLQCSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLSVSQLQDSG 180  
 DB 121 LVFGLTRANSDTHLLQCSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLSVSQLQDSG 180  
 QY 181 TWTCVTVLQONKKVEFKIDIVLAFOKAS 208  
 DB 181 TWTCVTVLQONKKVEFKIDIVLAFOKAS 208

RESULT 12  
 ADE57489  
 ID ADE57489 standard; protein; 458 AA.  
 XX  
 AC ADE57489;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human Protein P01730, SEQ ID NO 3351.  
 XX  
 KW Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SN1; Chung.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003016475-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 14-AUG-2002; 2002WO-US025765.  
 XX  
 PR 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-033347P.  
 XX  
 PA (GENO) GEN HOSPITAL CORP.  
 PA (PARB) BAYER AG.  
 XX  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX  
 DR WPI, 2003-268312/26.  
 DR GENBANK; P01730.

PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX  
 PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX

SQ Sequence 458 AA;

Query Match 39.9%; Score 1064; DB 7; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLLVQLALPAATQGGKVVLGKKGDVVELTCTASQKKSIOFHKNSNQIK 60  
 DB 1 MNRGVPFRHLVLLVQLALPAATQGGKVVLGKKGDVVELTCTASQKKSIOFHKNSNQIK 60  
 QY 61 ILNGQSSFLTKGPKSKLNDRAISRSLMDQGNFPLIKNLIKIEDSDTYICEVEDQKEEVOL 120  
 DB 61 ILNGQSSFLTKGPKSKLNDRAISRSLMDQGNFPLIKNLIKIEDSDTYICEVEDQKEEVOL 120  
 QY 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIOGKTLVSQLELQDSG 180  
 DB 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIOGKTLVSQLELQDSG 180  
 QY 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIOGKTLVSQLELQDSG 180  
 DB 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIOGKTLVSQLELQDSG 180  
 QY 181 TWTCTVLOQOKKVEFKIDIVLAFOKAS 208  
 DB 181 TWTCTVLOQOKKVEFKIDIVLAFOKAS 208

RESULT 13

ADP12495  
 ID ADP12495 standard; protein; 458 AA.

AC ADP12495;

DT 12-AUG-2004 (first entry)

DE Protein encoded by mRNA of the invention #105.

XX transplant rejection; immune system; rheumatoid arthritis; lupus;  
 KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS.

XX Homo sapiens.

XX WO2004042346-A2.  
 XX

PD 21-MAY-2004.

XX 24-APR-2003; 2003WO-US012946.

XX 24-APR-2002; 2002US-00131831.

XX 20-DEC-2002; 2002US-00325899.

PA (EXPR-) EXPRESSION DIAGNOSTICS INC.

XX Wohlgenuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;  
 PI Rosenbergs S;

DR WPI; 2004-400724/37.

XX Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,  
 PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant  
 PT rejection, in an individual, comprises detecting the expression level of  
 PT the genes.

PS Claim 65; SEQ ID NO 2504; 1762pp; English.

XX The present invention relates to diagnosing or monitoring transplant  
 CC rejection, e.g. cardiac or kidney transplant rejection, in an individual  
 CC comprising detecting the expression level of one or more genes. The  
 CC methods, system and kits are useful in diagnosing or monitoring  
 CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic  
 CC islet, lung, bone marrow or stem cell transplant rejection,  
 CC xenotransplant rejection or mechanical organ replacement rejection, in an  
 CC individual. The method is also useful in assessing the immune status of  
 CC an individual. The methods are also useful in diagnosing and monitoring  
 CC diseases that involve the immune system, e.g. rheumatoid arthritis,  
 CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or  
 CC viral, bacterial or fungal infection. The present sequence represents a  
 CC protein that is encoded by the mRNA of the invention.

SQ Sequence 458 AA;

Query Match 39.9%; Score 1064; DB 8; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLLVQLALPAATQGGKVVLGKKGDVVELTCTASQKKSIOFHKNSNQIK 60  
 DB 1 MNRGVPFRHLVLLVQLALPAATQGGKVVLGKKGDVVELTCTASQKKSIOFHKNSNQIK 60  
 QY 61 ILNGQSSFLTKGPKSKLNDRAISRSLMDQGNFPLIKNLIKIEDSDTYICEVEDQKEEVOL 120  
 DB 61 ILNGQSSFLTKGPKSKLNDRAISRSLMDQGNFPLIKNLIKIEDSDTYICEVEDQKEEVOL 120  
 QY 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIOGKTLVSQLELQDSG 180  
 DB 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIOGKTLVSQLELQDSG 180  
 QY 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIOGKTLVSQLELQDSG 180  
 DB 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIOGKTLVSQLELQDSG 180  
 QY 181 TWTCTVLOQOKKVEFKIDIVLAFOKAS 208  
 DB 181 TWTCTVLOQOKKVEFKIDIVLAFOKAS 208

RESULT 14

ADA44807  
 ID ADA44807 standard; protein; 473 AA.

AC ADA44807;

DT 04-DEC-2003 (first entry)

DE CD4/TCR CD3epsilon chain chimeric protein CD4epsilon15, SEQ ID NO:2.

XX HIV-1 infection; human immunodeficiency virus-1; CD4+ cell; chimeric CD4;  
 KW endoplasmic reticulum; ER retention; envelope protein gp160;  
 KW T cell receptor CD3epsilon chain; C-terminal domain; CD4epsilon15;  
 KW gene therapy; human; receptor.

XX

OS Chimeric.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Protein 1..458  
FT /label= CD4  
FT 459..473  
FT Region /note= "Part of the C-terminal domain of the T cell  
FT receptor CD3epsilon chain"  
XX  
XX MO2003076468-A1.  
XX  
XX 18-SEP-2003.  
XX  
XX 14-MAR-2003; 2003WO-ES000120.  
XX  
XX 14-MAR-2002; 2002ES-00000616.  
XX  
XX (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
XX  
XX Alarcon Sanchez BJ, San Jose Martinez ME, Zaldivar Notario I;  
PI Gomez Buendia M;  
XX  
XX WPI; 2003-779059/73.  
DR N-PSDB; ADA44806.  
XX  
XX  
XX Composition for treating or preventing human immune deficiency virus,  
PT comprises CD4 chimeric protein having a protective effect in trans, or  
PT related nucleic acid.  
XX  
XX Claim 5; Page 33-35; 43pp; Spanish.  
XX  
XX The invention relates to a composition for the treatment or prevention of  
CC human immunodeficiency virus-1 (HIV-1) infection. The composition  
CC comprises CD4+ cells that have been transduced with a vector that encodes  
CC a chimeric CD4 molecule which is capable of being retained in the  
CC endoplasmic reticulum (ER). The invention also encompasses the use of a  
CC soluble protein factor produced by CD4+ cells that have been transduced  
CC with a vector encoding a chimeric CD4 protein; and the use of an  
CC expression system encoding a chimeric CD4 protein. The ER-localised  
CC chimeric CD4 molecule binds to the HIV-1 envelope protein gp160,  
CC resulting in HIV-1 retention in the ER and thereby preventing viral  
CC replication. In a specific embodiment, the chimeric CD4 molecule  
CC comprises CD4 fused to 15 amino acids of the C-terminal domain of the T  
CC cell receptor CD3epsilon chain; this chimeric CD4 molecule is designated  
CC CD4epsilon15 (ADA44807). A known chimeric CD4 of similar structure but  
CC containing only 10 amino acids from CD3epsilon can also be used.  
CC Compositions of the invention have an in trans effect on the replication  
CC of HIV-1, and may be used to treat and prevent HIV-1 infection. The  
CC present sequence represents the chimeric CD4 molecule CD4epsilon15, which  
CC is specifically claimed for use in compositions of the invention.  
XX  
XX  
SQ Sequence 473 AA;  
Query Match 39.9%; Score 1064; DB 7; Length 473;  
Best Local Similarity 100.0%; Pred. No. 9, 8e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 15  
AAV59170  
ID AAV59170 standard; protein; 474 AA.  
XX  
XX AAV59170;  
AC  
XX  
XX 14-MAR-2000 (first entry)  
DT  
XX  
XX CD4-Ig fusion protein CD4Mmu.  
DE  
XX  
XX HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;  
XX secreted protein; SIV infection; medicament.  
KM  
XX  
XX Synthetic.  
OS  
XX Homo sapiens.  
XX  
XX CA1340741-C.  
PN  
XX  
XX 14-SEP-1999.  
PD  
XX  
XX 20-JAN-1989; 89CA-00588749.  
PF  
XX  
XX 20-JAN-1989; 89CA-00588749.  
PR  
XX  
XX (GENO ) GEN HOSPITAL CORP.  
PA  
XX  
XX Seed B;  
PI  
XX  
XX WPI; 2000-063015/06.  
DR N-PSDB; AA248203.  
XX  
XX  
XX The invention provides a fusion gene encoding a fusion protein that  
CC comprises an extracellular CD4 DNA sequence or its fragment which binds  
CC to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA  
CC sequence of an Ig heavy or light chain, where the DNA sequence encoding  
CC the variable region has been replaced with the DNA sequence which encodes  
CC extracellular CD4 or its gp120 binding fragment. The fusion protein is  
CC capable of being secreted. The fusion proteins are useful for treating  
CC HIV or SIV infections in animals, preferably humans. They are also useful  
CC for producing medicaments which can be used for treating HIV or SIV  
CC infections in humans. The present sequence represents the fusion protein  
CC CD4Mmu where the CD4 is linked to human IgG1 at the Met2 site upstream of  
CC the CH1 region  
XX  
XX  
SQ Sequence 474 AA;  
Query Match 39.9%; Score 1064; DB 3; Length 474;  
Best Local Similarity 100.0%; Pred. No. 9, 8e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: September 20, 2005, 16:07:08  
Job time : 121 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

## OM protein - protein search, using sw model

Run on: September 20, 2005, 15:24:08 ; Search time 32 Seconds  
(without alignments)  
1185.054 Million cell updates/sec

Title: US-09-936-702-3

Perfect score: 2668

Sequence: 1 MNRGVPFRHLVLVQLALP.....GTRLEIKLVPRGSGHHHHH 508

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*

2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*

3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*

4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*

6: /cgn2\_6/ptodata/1/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Query length	ID	Description
1	1064	39.9	295	6 5223394-9	Patent No. 5223394
2	1064	39.9	295	6 5223394-9	Patent No. 5223394
3	1064	39.9	318	6 5223394-11	Patent No. 5223394
4	1064	39.9	318	6 5223394-11	Patent No. 5223394
5	1064	39.9	458	3 US-09-039-555B-15	Sequence 15, Appl
6	1064	39.9	458	3 US-09-517-605-3	Sequence 3, Appl
7	1062	39.8	458	4 US-10-092-138A-25	Sequence 25, Appl
8	1059	39.7	398	4 US-08-284-391B-29	Sequence 29, Appl
9	1059	39.7	398	4 US-09-218-950-29	Sequence 29, Appl
10	1059	39.7	398	4 US-08-394-388A-29	Sequence 29, Appl
11	1059	39.7	402	1 US-08-236-311-1	Sequence 1, Appl
12	1059	39.7	402	1 US-08-457-918-1	Sequence 1, Appl
13	1059	39.7	402	4 US-10-157-408-1	Sequence 1, Appl
14	1059	39.7	457	4 US-08-328-500-9	Sequence 9, Appl
15	1059	39.7	458	3 US-08-466-368-4	Sequence 4, Appl
16	1059	39.7	462	2 US-08-417-495-5	Sequence 5, Appl
17	1059	39.7	462	2 US-08-284-391B-5	Sequence 5, Appl
18	1059	39.7	462	3 US-09-218-950-5	Sequence 5, Appl
19	1059	39.7	462	4 US-08-394-388A-5	Sequence 5, Appl
20	1059	39.7	462	5 PCT-US92-01785-5	Sequence 5, Appl
21	1059	39.7	462	5 PCT-US95-00454-5	Sequence 5, Appl
22	1059	39.7	532	2 US-08-417-495-6	Sequence 6, Appl
23	1059	39.7	532	2 US-08-284-391B-6	Sequence 6, Appl
24	1059	39.7	532	2 US-09-218-950-6	Sequence 6, Appl
25	1059	39.7	532	4 US-08-394-388A-6	Sequence 6, Appl
26	1059	39.7	532	4 PCT-US92-01785-6	Sequence 6, Appl
27	1059	39.7	532	5 PCT-US95-00454-6	Sequence 6, Appl

28	1059	39.7	575	2 US-08-417-495-4	Sequence 4, Appl
29	1059	39.7	575	2 US-08-284-391B-4	Sequence 4, Appl
30	1059	39.7	575	3 US-09-218-950-4	Sequence 4, Appl
31	1059	39.7	575	4 US-08-394-388A-4	Sequence 4, Appl
32	1059	39.7	575	5 PCT-US92-01785-4	Sequence 4, Appl
33	1059	39.7	575	5 PCT-US95-00454-4	Sequence 4, Appl
34	1059	39.7	630	4 US-08-472-888A-6	Sequence 6, Appl
35	1059	39.6	458	6 5223394-7	Patent No. 5223394
36	1056	39.6	458	6 5223394-7	Patent No. 5223394
37	1053	39.5	394	4 US-08-466-368-2	Sequence 2, Appl
38	1053	39.5	394	4 US-08-328-500-2	Sequence 2, Appl
39	1051	39.4	530	3 US-08-477-460B-4	Sequence 4, Appl
40	1051	39.4	530	3 US-08-379-516-4	Sequence 4, Appl
41	1051	39.4	530	3 US-09-329-916-4	Sequence 4, Appl
42	1051	39.4	530	3 US-08-485-372A-4	Sequence 4, Appl
43	1051	39.4	530	3 US-09-409-006A-4	Sequence 4, Appl
44	1051	39.4	530	4 US-08-484-681-4	Sequence 4, Appl
45	1051	39.4	530	4 US-09-766-995-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1

5223394-9

Patent No. 5223394

APPLICANT: WALLNER, BARBARA

TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING

LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL

LINKAGE SIGNAL SEQUENCE

NUMBER OF SEQUENCES: 12

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/335,688

FILING DATE: 10-APR-1989

SEQ ID NO:9:

LENGTH: 295

5223394-9

Query Match 39.9%; Score 1064; DB 6; Length 295;

Best Local Similarity 100.0%; Pred. No. 9.1e-54;

Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPAAATGKKVVLGGKGDIVELTCTASQKKSIOFHAKNSNQIK 60

DB 1 MNRGVPFRHLVLVQLALPAAATGKKVVLGGKGDIVELTCTASQKKSIOFHAKNSNQIK 60

QY 61 ILGNQGSFLTKGSPSKLNDRAISRSLMDQGNFLLIKNLKIEDSDTYICEVEQKEVQL 120

DB 61 ILGNQGSFLTKGSPSKLNDRAISRSLMDQGNFLLIKNLKIEDSDTYICEVEQKEVQL 120

QY 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPVQCRSPGKNTQGGKTLSVSQLELDPSG 180

DB 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPVQCRSPGKNTQGGKTLSVSQLELDPSG 180

QY 181 TWTCTVLQNKQKVEFKIDIVLAFQKAS 208

DB 181 TWTCTVLQNKQKVEFKIDIVLAFQKAS 208

RESULT 2

5223394-9

Patent No. 5223394

APPLICANT: WALLNER, BARBARA

TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING

LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL

LINKAGE SIGNAL SEQUENCE

NUMBER OF SEQUENCES: 12

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/335,688

FILING DATE: 10-APR-1989

SEQ ID NO:9:

LENGTH: 295

5223394-9

Query Match 39.9%; Score 1064; DB 6; Length 295;  
Best Local Similarity 100.0%; Pred. No. 9.1e-54;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLILVQLALLPAATGKKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60  
DB 1 MNRGVPFRHLILVQLALLPAATGKKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60  
QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFLLIKNLKIBSDTYICEVEDQKEEVL 120  
DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFLLIKNLKIBSDTYICEVEDQKEEVL 120  
QY 121 LVFGLTANSDTHLLQGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVQLQLDQSG 180  
DB 121 LVFGLTANSDTHLLQGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVQLQLDQSG 180  
QY 181 TWCTCVLQONOKVFEKIDIVLAFQKAS 208  
DB 181 TWCTCVLQONOKVFEKIDIVLAFQKAS 208

RESULT 3  
5223394-11

APPLICANT: WALLNER, BARBARA  
TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING  
LVMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL  
LINKAGE SIGNAL SEQUENCE  
NUMBER OF SEQUENCES: 12  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/335,688  
FILING DATE: 10-APR-1989  
SEQ ID NO:11:  
LENGTH: 318

5223394-11

Query Match 39.9%; Score 1064; DB 6; Length 318;  
Best Local Similarity 100.0%; Pred. No. 9.8e-54;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLILVQLALLPAATGKKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60  
DB 1 MNRGVPFRHLILVQLALLPAATGKKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60  
QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFLLIKNLKIBSDTYICEVEDQKEEVL 120  
DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFLLIKNLKIBSDTYICEVEDQKEEVL 120  
QY 121 LVFGLTANSDTHLLQGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVQLQLDQSG 180  
DB 121 LVFGLTANSDTHLLQGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVQLQLDQSG 180  
QY 181 TWCTCVLQONOKVFEKIDIVLAFQKAS 208  
DB 181 TWCTCVLQONOKVFEKIDIVLAFQKAS 208

RESULT 4  
5223394-11

APPLICANT: WALLNER, BARBARA  
TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING  
LVMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL  
LINKAGE SIGNAL SEQUENCE  
NUMBER OF SEQUENCES: 12  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/335,688  
FILING DATE: 10-APR-1989  
SEQ ID NO:11:  
LENGTH: 318

5223394-11

Query Match 39.9%; Score 1064; DB 6; Length 318;  
Best Local Similarity 100.0%; Pred. No. 9.8e-54;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLILVQLALLPAATGKKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60  
DB 1 MNRGVPFRHLILVQLALLPAATGKKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60  
QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFLLIKNLKIBSDTYICEVEDQKEEVL 120  
DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFLLIKNLKIBSDTYICEVEDQKEEVL 120  
QY 121 LVFGLTANSDTHLLQGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVQLQLDQSG 180  
DB 121 LVFGLTANSDTHLLQGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVQLQLDQSG 180  
QY 181 TWCTCVLQONOKVFEKIDIVLAFQKAS 208  
DB 181 TWCTCVLQONOKVFEKIDIVLAFQKAS 208

RESULT 5

US-09-039-555B-15  
Sequence 15, Application US/09039555B

PATENT No. 6033856  
GENERAL INFORMATION:  
APPLICANT: Koerner, Kathrin  
APPLICANT: Mueller, Rolf  
TITLE OF INVENTION: PROMOTER OF THE CDC25B GENE, ITS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/039,555B  
FILING DATE: 16-MAR-1998  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 19710643.9  
FILING DATE: 14-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Bent, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 016779/0131  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 458 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-039-555B-15

Query Match 39.9%; Score 1064; DB 3; Length 458;  
Best Local Similarity 100.0%; Pred. No. 1.4e-53;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLILVQLALLPAATGKKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60



```
Db 1 MNRGVPFRHLLVQLALLPAATQGGKVVLTCTASQKKSIOFHWKNSNOIK 60
Qy 61 ILGNQGSFLTKGSPKNDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120
Db 61 ILGNQGSFLTKGSPKNDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120
Qy 121 LVFGLTANSDTHLLQGSGLTITLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db 121 LVFGLTANSDTHLLQGSGLTITLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy 181 TWCTVLOQNKVFEFKIDIVLAFQKAS 208
Db 181 TWCTVLOQNKVFEFKIDIVLAFQKAS 208

RESULT 6
US-09-517-605-3
; Sequence 3, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Littenman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geltenbeck, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-517-605-3

Query Match 39.9%; Score 1064; DB 3; Length 458;
Best Local Similarity 100.0%; Pred. No. 1,4e-53;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNRGVPFRHLLVQLALLPAATQGGKVVLTCTASQKKSIOFHWKNSNOIK 60
Db 1 MNRGVPFRHLLVQLALLPAATQGGKVVLTCTASQKKSIOFHWKNSNOIK 60
Qy 61 ILGNQGSFLTKGSPKNDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120
Db 61 ILGNQGSFLTKGSPKNDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120
Qy 121 LVFGLTANSDTHLLQGSGLTITLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db 121 LVFGLTANSDTHLLQGSGLTITLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy 181 TWCTVLOQNKVFEFKIDIVLAFQKAS 208
Db 181 TWCTVLOQNKVFEFKIDIVLAFQKAS 208

RESULT 7
US-10-092-138A-25
; Sequence 25, Application US/10092138A
; Patent No. 6743530
; GENERAL INFORMATION:
; APPLICANT: Sato, Taka-Aki
; TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
; FILE REFERENCE: 65823/JPW/PT
; CURRENT APPLICATION NUMBER: US/10/092,138A
; CURRENT FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 25
```

```
; LENGTH: 458
; TYPE: PRT
; ORGANISM: human
US-10-092-138A-25

Query Match 39.8%; Score 1062; DB 4; Length 458;
Best Local Similarity 99.5%; Pred. No. 1.8e-53;
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNRGVPFRHLLVQLALLPAATQGGKVVLTCTASQKKSIOFHWKNSNOIK 60
Db 1 MNRGVPFRHLLVQLALLPAATQGGKVVLTCTASQKKSIOFHWKNSNOIK 60
Qy 61 ILGNQGSFLTKGSPKNDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120
Db 61 ILGNQGSFLTKGSPKNDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120
Qy 121 LVFGLTANSDTHLLQGSGLTITLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db 121 LVFGLTANSDTHLLQGSGLTITLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy 181 TWCTVLOQNKVFEFKIDIVLAFQKAS 208
Db 181 TWCTVLOQNKVFEFKIDIVLAFQKAS 208

RESULT 8
US-08-284-391B-29
; Sequence 29, Application US/08284391B
; Patent No. 5851828
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391B
; FILING DATE: 02-AUG-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
```

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-284-391B-29

Query Match 39.7%; Score 1059; DB 2; Length 398;  
Best Local Similarity 99.5%; Pred. No. 2,3e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALLPAATQGGKVVLTGKGDVVELTCTASQKKSIOFHWKNSNQIK 60  
DB 1 MNRGVPFRHLVLVQLALLPAATQGGKVVLTGKGDVVELTCTASQKKSIOFHWKNSNQIK 60  
QY 61 ILNQGSPFLTKGSPSKINDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILNQGSPFLTKGSPSKINDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLOQGSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLSVSQLELDGSG 180  
DB 121 LVFGLTANSDTHLLOQGSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLSVSQLELDGSG 180  
QY 181 TWTCTVLQONQKVEFKIDIVLAFOKAS 208  
DB 181 TWTCTVLQONQKVEFKIDIVLAFOKAS 208

## RESULT 9

US-09-218-950-29  
Sequence 29, Application US/09218950  
Patent No. 6284240  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/218,950  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:

LENGTH: 398 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-218-950-29

Query Match 39.7%; Score 1059; DB 3; Length 398;  
Best Local Similarity 99.5%; Pred. No. 2,3e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALLPAATQGGKVVLTGKGDVVELTCTASQKKSIOFHWKNSNQIK 60  
DB 1 MNRGVPFRHLVLVQLALLPAATQGGKVVLTGKGDVVELTCTASQKKSIOFHWKNSNQIK 60  
QY 61 ILNQGSPFLTKGSPSKINDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILNQGSPFLTKGSPSKINDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLOQGSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLSVSQLELDGSG 180  
DB 121 LVFGLTANSDTHLLOQGSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLSVSQLELDGSG 180  
QY 181 TWTCTVLQONQKVEFKIDIVLAFOKAS 208  
DB 181 TWTCTVLQONQKVEFKIDIVLAFOKAS 208

## RESULT 10

US-08-394-388A-29  
Sequence 29, Application US/08394388A  
Patent No. 6753162  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/394,388A  
FILING DATE: 24-FEB-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:

INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 398 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-394-388A-29

Query Match 39.7%; Score 1059; DB 4; Length 398;  
Best Local Similarity 99.5%; Pred. No. 2,36-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALPAATQKKVYLGKGDVTELTCTASQKSIQFHKNSNQIK 60  
DB 1 MNRGVPFRHLLVQLALPAATQKKVYLGKGDVTELTCTASQKSIQFHKNSNQIK 60  
QY 61 ILGNQGSFLTQKPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLTQKPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGSLLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLQGSLLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWICTVLQNKQKVEFKIDIVVLAFOKAS 208  
DB 181 TWICTVLQNKQKVEFKIDIVVLAFOKAS 208

## RESULT 11

US-08-236-311-1  
Sequence 1, Application US/08236311

PATENT NO. 5565335  
GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
APPLICANT: Gregory, Timothy J.  
TITLE OF INVENTION: Adhesion Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 KB floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/236,311  
FILING DATE: 02-MAY-1994  
CLASSIFICATION: 435  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Haack, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 444PIC2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896

TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-236-311-1

Query Match 39.7%; Score 1059; DB 1; Length 402;  
Best Local Similarity 99.5%; Pred. No. 2,46-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALPAATQKKVYLGKGDVTELTCTASQKSIQFHKNSNQIK 60  
DB 1 MNRGVPFRHLLVQLALPAATQKKVYLGKGDVTELTCTASQKSIQFHKNSNQIK 60  
QY 61 ILGNQGSFLTQKPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLTQKPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGSLLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLQGSLLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWICTVLQNKQKVEFKIDIVVLAFOKAS 208  
DB 181 TWICTVLQNKQKVEFKIDIVVLAFOKAS 208

## RESULT 12

US-08-457-918-1  
Sequence 1, Application US/08457918

PATENT NO. 617655  
GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
APPLICANT: Gregory, Timothy J.  
TITLE OF INVENTION: Adhesion Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 KB floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,918  
FILING DATE: 1-JUN-1995  
CLASSIFICATION: 435  
APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Kubinec, Jeffrey S.  
REGISTRATION NUMBER: 36,575

REFERENCE/DOCKET NUMBER: P0444PIC3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-457-918-1

Query Match 39.7%; Score 1059; DB 3; Length 402;  
Best Local Similarity 99.5%; Pred. No. 2,4e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALLPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHKNSNOIK 60  
DB 1 MNRGVPFRHLVLVQLALLPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHKNSNOIK 60  
QY 61 ILNGQSFLLTKGSKLNDRADSRSLMDQGNFPLIIKLIKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILNGQSFLLTKGSKLNDRADSRSLMDQGNFPLIIKLIKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSPDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSPDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLQONOKVFEKIDIVLAFQKAS 208  
DB 181 TWTCTVLQONOKVFEKIDIVLAFQKAS 208

## RESULT 13

US-10-157-408-1  
Sequence 1, Application US/10157408  
Patent No. 6710169  
GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
Gregory, Timothy J.  
TITLE OF INVENTION: Adhesion Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/157,408  
FILING DATE: 28-May-2002  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,918  
FILING DATE: 1-JUN-1995  
APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Kubinec, Jeffrey S.

REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0444PIC3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-157-408-1

Query Match 39.7%; Score 1059; DB 4; Length 402;  
Best Local Similarity 99.5%; Pred. No. 2,4e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALLPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHKNSNOIK 60  
DB 1 MNRGVPFRHLVLVQLALLPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHKNSNOIK 60  
QY 61 ILNGQSFLLTKGSKLNDRADSRSLMDQGNFPLIIKLIKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILNGQSFLLTKGSKLNDRADSRSLMDQGNFPLIIKLIKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSPDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSPDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLQONOKVFEKIDIVLAFQKAS 208  
DB 181 TWTCTVLQONOKVFEKIDIVLAFQKAS 208

## RESULT 14

US-08-328-500-9  
Sequence 9, Application US/08328500  
Patent No. 6673896  
GENERAL INFORMATION:  
APPLICANT: Madden, Paul J.  
APPICANT: Axel, Richard  
APPICANT: Sweet, Richard W.  
APPICANT: Archos, James  
TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/328,500  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0575/24577-CY  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 457 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-328-500-9

Query Match 39.7%; Score 1059; DB 4; Length 457;  
Best Local Similarity 99.5%; Pred. No. 2.7e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALPPATOGKKVVLGKKGDTVELTCTASQKSIQFHKNSNQIK 60  
DB 1 MNRGVPRHLLVQLALPPATOGKKVVLGKKGDTVELTCTASQKSIQFHKNSNQIK 60  
QY 61 ILGNQGSFLTKGPKSKINDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLTKGPKSKINDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSPTHLQGSQSLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSPTHLQGSQSLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLQNKQKVEFKIDIVLAFOKAS 208  
DB 181 TWTCTVLQNKQKVEFKIDIVLAFOKAS 208

## RESULT 15

US-08-466-368-4  
Sequence 4, Application US/08466368  
Patent No. 6093539

## GENERAL INFORMATION:

APPLICANT: Maddon, Paul J.  
APPLICANT: Littman, Dan R.  
APPLICANT: Chess, Leonard  
APPLICANT: Axel, Richard  
APPLICANT: Weiss, Robin  
APPLICANT: McDougall, J. S.  
TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN  
TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

## COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,368

FILING DATE: 06-JUN-1995

## CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 24577-E1-B/JPM/AKC

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-391-0525

TELEFAX: 212-278-0400

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 458 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-466-368-4

QY 1 MNRGVPRHLLVQLALPPATOGKKVVLGKKGDTVELTCTASQKSIQFHKNSNQIK 60  
DB 1 MNRGVPRHLLVQLALPPATOGKKVVLGKKGDTVELTCTASQKSIQFHKNSNQIK 60  
QY 61 ILGNQGSFLTKGPKSKINDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLTKGPKSKINDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSPTHLQGSQSLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSPTHLQGSQSLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLQNKQKVEFKIDIVLAFOKAS 208  
DB 181 TWTCTVLQNKQKVEFKIDIVLAFOKAS 208

Search completed: September 20, 2005, 16:10:48  
Job time: 34 secs

Query Match 39.7%; Score 1059; DB 3; Length 458;  
Best Local Similarity 99.5%; Pred. No. 2.7e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 16:09:38 ; Search time 485 Seconds  
(without alignments)  
424.130 Million cell updates/sec

Title: US-09-936-702-3  
Perfect score: 2668  
Sequence: 1 MNRGVPFRHLLVLQALALP.....GTRLEIKLVPRSGSHHHHH 508

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1812044 seqs, 404927589 residues

Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgnt\_6/prodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgnt\_6/prodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*  
3: /cgnt\_6/prodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgnt\_6/prodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgnt\_6/prodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
6: /cgnt\_6/prodata/1/pubppaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgnt\_6/prodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
8: /cgnt\_6/prodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgnt\_6/prodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
10: /cgnt\_6/prodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgnt\_6/prodata/1/pubppaa/US09C\_PUBCOMB.pep:\*  
12: /cgnt\_6/prodata/1/pubppaa/US09C\_NEW\_PUB.pep:\*  
13: /cgnt\_6/prodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
14: /cgnt\_6/prodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
15: /cgnt\_6/prodata/1/pubppaa/US10C\_PUBCOMB.pep:\*  
16: /cgnt\_6/prodata/1/pubppaa/US10C\_PUBCOMB.pep:\*  
17: /cgnt\_6/prodata/1/pubppaa/US10E\_PUBCOMB.pep:\*  
18: /cgnt\_6/prodata/1/pubppaa/US10E\_NEW\_PUB.pep:\*  
19: /cgnt\_6/prodata/1/pubppaa/US11\_PUBCOMB.pep:\*  
20: /cgnt\_6/prodata/1/pubppaa/US11\_NEW\_PUB.pep:\*  
21: /cgnt\_6/prodata/1/pubppaa/US60\_NEW\_PUB.pep:\*  
22: /cgnt\_6/prodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1064	39.9	458	US-10-151-274-3	Sequence 3, App1
2	1064	39.9	458	US-10-103-597A-39	Sequence 39, App1
3	1064	39.9	458	US-10-188-444-39	Sequence 39, App1
4	1064	39.9	458	US-10-207-655-170	Sequence 170, App1
5	1062	39.8	458	US-08-681-219-27	Sequence 27, App1
6	1062	39.8	458	US-09-230-111C-25	Sequence 25, App1
7	1062	39.8	458	US-10-092-138-25	Sequence 25, App1
8	1062	39.8	458	US-10-820-403-25	Sequence 25, App1
9	1059	39.7	458	US-09-939-537-29	Sequence 29, App1
10	1059	39.7	402	US-10-157-408-1	Sequence 1, App1
11	1059	39.7	402	US-10-097-044A-1	Sequence 1, App1

12	1059	39.7	402	US-10-769-247-1	Sequence 1, App1
13	1059	39.7	457	US-09-891-119A-9	Sequence 9, App1
14	1059	39.7	462	US-09-939-537-5	Sequence 5, App1
15	1059	39.7	462	US-09-243-008-5	Sequence 5, App1
16	1059	39.7	532	US-09-939-537-6	Sequence 6, App1
17	1059	39.7	532	US-09-243-008-6	Sequence 6, App1
18	1059	39.7	575	US-09-939-537-4	Sequence 4, App1
19	1059	39.7	575	US-09-243-008-4	Sequence 4, App1
20	1053	39.5	397	US-09-891-119A-2	Sequence 2, App1
21	1051	39.4	530	US-08-485-163-5	Sequence 5, App1
22	1051	39.4	530	US-09-766-995-4	Sequence 4, App1
23	1050	39.4	432	US-08-485-163-3	Sequence 3, App1
24	1050	39.4	432	US-09-766-995-2	Sequence 2, App1
25	1048	39.3	310	US-08-485-163-7	Sequence 7, App1
26	1048	39.3	310	US-09-766-995-6	Sequence 6, App1
27	1040	39.0	203	US-10-513-725-10	Sequence 10, App1
28	1040	39.0	203	US-10-512-966-7	Sequence 7, App1
29	1040	39.0	449	US-10-493-676-11	Sequence 11, App1
30	1040	39.0	450	US-10-493-676-6	Sequence 6, App1
31	1040	39.0	487	US-10-512-966-8	Sequence 8, App1
32	1036	38.8	202	US-10-492-729-19	Sequence 19, App1
33	1035	38.8	203	US-09-939-537-31	Sequence 31, App1
34	959.5	36.0	250	US-09-880-748-1952	Sequence 1952, App
35	959.5	36.0	250	US-10-293-418-1952	Sequence 1952, App
36	957.5	35.9	246	US-09-880-748-1268	Sequence 1268, App
37	957.5	35.9	246	US-10-293-418-1268	Sequence 1268, App
38	954	35.8	251	US-09-880-748-1238	Sequence 1238, App
39	954	35.8	251	US-10-293-418-1238	Sequence 1238, App
40	946	35.5	434	US-10-157-408-4	Sequence 4, App1
41	946	35.5	434	US-10-097-044A-4	Sequence 4, App1
42	946	35.5	444	US-10-769-247-4	Sequence 4, App1
43	945.5	35.4	448	US-10-024-329-32	Sequence 32, App1
44	940	35.2	370	US-09-759-841-6	Sequence 6, App1
45	940	35.2	433	US-10-872-198-129	Sequence 129, App

ALIGNMENTS

RESULT 1  
US-10-151-274-3  
; Sequence 3, Application US/10151274  
; Publication No. US20030064071A1  
GENERAL INFORMATION:  
APPLICANT: Littman, Dan R.  
APPLICANT: Kwon, Douglas S.  
APPLICANT: van Kooyk, Yvette  
APPLICANT: Geiljenbeck, Theo  
TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY  
TITLE OF INVENTION: INVO  
FILE REFERENCE: 1049-1-017  
CURRENT APPLICATION NUMBER: US/10/151,274  
CURRENT FILING DATE: 2002-05-20  
PRIOR APPLICATION NUMBER: US/09/517,605  
PRIOR FILING DATE: 2000-03-02  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 458  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-151-274-3  
Query Match 39.9%; Score 1064; DB 14; Length 458;  
Best Local Similarity 100.0%; Pred. No. 7.4e-51;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNRGVPFRHLLVLQALALPAAATGGKVVYLTCTASOKKSIOFHKNNOIK 60  
DB 1 MNRGVPFRHLLVLQALALPAAATGGKVVYLTCTASOKKSIOFHKNNOIK 60  
QY 61 ILNQGSLFKGPKSLNDRADSRSLMDQGNFPLIIKNLIKEDSDTYICEVEDOKEEVOL 120

```
Db      61  |||||
        ILGNQGSFLTKGSPSKLNDRAADRSRLWDQGNFPLIKNLKIEBSDTYICEVEDQKEEVOL 120
Qy      121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
        |||||
Db      121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy      181 TWCTCTVLQONQKVEFKIDIVVLAFOKAS 208
        |||||
Db      181 TWCTCTVLQONQKVEFKIDIVVLAFOKAS 208
```

```
RESULT 2
US-10-103-597A-39
; Sequence 39, Application US/10103597A
; Publication No. US20030096432A1
; GENERAL INFORMATION:
; APPLICANT: Jakobsen, Bent Karsten
; TITLE OF INVENTION: Screening Methods
; FILE REFERENCE: 102286.142
; CURRENT APPLICATION NUMBER: US/10/103,597A
; PRIOR FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: PCT/GB00/03579
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 9922352.1
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-103-597A-39

Query Match      39.9%; Score 1064; DB 14; Length 458;
Best Local Similarity 100.0%; Pred. No. 7.4e-51;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MNRGVPRHLLVLQALPLPAATQGGKVVLGKKGDTVELTCTASQKKSIQFHWKNSNOIK 60
Db      1  MNRGVPRHLLVLQALPLPAATQGGKVVLGKKGDTVELTCTASQKKSIQFHWKNSNOIK 60
Qy      61  ILGNQGSFLTKGSPSKLNDRAADRSRLWDQGNFPLIKNLKIEBSDTYICEVEDQKEEVOL 120
Db      61  ILGNQGSFLTKGSPSKLNDRAADRSRLWDQGNFPLIKNLKIEBSDTYICEVEDQKEEVOL 120
Qy      121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db      121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy      181 TWCTCTVLQONQKVEFKIDIVVLAFOKAS 208
Db      181 TWCTCTVLQONQKVEFKIDIVVLAFOKAS 208

RESULT 3
US-10-188-444-39
; Sequence 39, Application US/10188444
; Publication No. US20030104635A1
; GENERAL INFORMATION:
; APPLICANT: Jakobsen, Bent Karsten
; TITLE OF INVENTION: Screening Methods
; FILE REFERENCE: 102286.142 (CIP)
; CURRENT APPLICATION NUMBER: US/10/188,444
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: PCT/GB00/03579
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 9922352.1
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 458
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-444-39

Query Match      39.9%; Score 1064; DB 14; Length 458;
Best Local Similarity 100.0%; Pred. No. 7.4e-51;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MNRGVPRHLLVLQALPLPAATQGGKVVLGKKGDTVELTCTASQKKSIQFHWKNSNOIK 60
Db      1  MNRGVPRHLLVLQALPLPAATQGGKVVLGKKGDTVELTCTASQKKSIQFHWKNSNOIK 60
Qy      61  ILGNQGSFLTKGSPSKLNDRAADRSRLWDQGNFPLIKNLKIEBSDTYICEVEDQKEEVOL 120
Db      61  ILGNQGSFLTKGSPSKLNDRAADRSRLWDQGNFPLIKNLKIEBSDTYICEVEDQKEEVOL 120
Qy      121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db      121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy      181 TWCTCTVLQONQKVEFKIDIVVLAFOKAS 208
Db      181 TWCTCTVLQONQKVEFKIDIVVLAFOKAS 208
```

```
RESULT 4
US-10-207-655-170
; Sequence 170, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; PRIOR FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 170
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-170

Query Match      39.9%; Score 1064; DB 14; Length 458;
Best Local Similarity 100.0%; Pred. No. 7.4e-51;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MNRGVPRHLLVLQALPLPAATQGGKVVLGKKGDTVELTCTASQKKSIQFHWKNSNOIK 60
Db      1  MNRGVPRHLLVLQALPLPAATQGGKVVLGKKGDTVELTCTASQKKSIQFHWKNSNOIK 60
Qy      61  ILGNQGSFLTKGSPSKLNDRAADRSRLWDQGNFPLIKNLKIEBSDTYICEVEDQKEEVOL 120
Db      61  ILGNQGSFLTKGSPSKLNDRAADRSRLWDQGNFPLIKNLKIEBSDTYICEVEDQKEEVOL 120
Qy      121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db      121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy      181 TWCTCTVLQONQKVEFKIDIVVLAFOKAS 208
Db      181 TWCTCTVLQONQKVEFKIDIVVLAFOKAS 208

RESULT 5
US-08-681-219-27
; Sequence 27, Application US/08681219
; Publication No. US20020058607A1
; GENERAL INFORMATION:
; APPLICANT: Takaaki Sato and Junn Yanagisawa
; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN
; SIGNAL-TRANSDUCING PROTEINS AND THE GPCR
```



```

; TITLE OF INVENTION: (PDZ/DHR) DOMAIN AND USES THEREOF
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/681,219
; FILING DATE: 22-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/48962/JPM/JKM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 458 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-681-219-27

```

```

Query Match      39.8%; Score 1062; DB 8; Length 458;
Best Local Similarity 99.5%; Pred. No. 9.5e-51;
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MNRGVPFRHLVLTQALIPAAATQGGKGVVGGKGDVVELCTASQKKSIOFHKNSNQIK 60
DB 1 MNRGVPFRHLVLTQALIPAAATQGGKGVVGGKGDVVELCTASQKKSIOFHKNSNQIK 60
QY 61 ILGNQGSFLTKGSPSKINDRADSRSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGSPSKINDRADSRSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGQSLTLTLESPPGSSPSVQCSPPGKNIOGKKTLSVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGQSLTLTLESPPGSSPSVQCSPPGKNIOGKKTLSVSQLELDQSG 180
QY 181 TWTCTVLQNGKVEFKIDIVLAFOKAS 208
DB 181 TWTCTVLQNGKVEFKIDIVLAFOKAS 208

```

```

RESULT 6
US-09-230-111C-25
; Sequence 25, Application US/09230111C
; Publication No. US20030203414A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Taka-Aki
; APPLICANT: Yanagisawa, Jun
; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN
; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF (PDZ/DHR)
; FILE REFERENCE: 48962-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,111C
; CURRENT FILING DATE: 1999-05-17
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 458
; TYPE: PRT

```

```

; ORGANISM: human
US-09-230-111C-25

```

```

Query Match      39.8%; Score 1062; DB 10; Length 458;
Best Local Similarity 99.5%; Pred. No. 9.5e-51;
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MNRGVPFRHLVLTQALIPAAATQGGKGVVGGKGDVVELCTASQKKSIOFHKNSNQIK 60
DB 1 MNRGVPFRHLVLTQALIPAAATQGGKGVVGGKGDVVELCTASQKKSIOFHKNSNQIK 60
QY 61 ILGNQGSFLTKGSPSKINDRADSRSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGSPSKINDRADSRSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGQSLTLTLESPPGSSPSVQCSPPGKNIOGKKTLSVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGQSLTLTLESPPGSSPSVQCSPPGKNIOGKKTLSVSQLELDQSG 180
QY 181 TWTCTVLQNGKVEFKIDIVLAFOKAS 208
DB 181 TWTCTVLQNGKVEFKIDIVLAFOKAS 208

```

```

RESULT 7
US-10-092-138-25
; Sequence 25, Application US/10092138
; Publication No. US20030170723A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Taka-Aki
; TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
; TITLE OF INVENTION: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: 65823/JPW/PT
; CURRENT APPLICATION NUMBER: US/10/092,138
; CURRENT FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 458
; TYPE: PRT
; ORGANISM: human
US-10-092-138-25

```

```

Query Match      39.8%; Score 1062; DB 14; Length 458;
Best Local Similarity 99.5%; Pred. No. 9.5e-51;
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MNRGVPFRHLVLTQALIPAAATQGGKGVVGGKGDVVELCTASQKKSIOFHKNSNQIK 60
DB 1 MNRGVPFRHLVLTQALIPAAATQGGKGVVGGKGDVVELCTASQKKSIOFHKNSNQIK 60
QY 61 ILGNQGSFLTKGSPSKINDRADSRSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGSPSKINDRADSRSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGQSLTLTLESPPGSSPSVQCSPPGKNIOGKKTLSVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGQSLTLTLESPPGSSPSVQCSPPGKNIOGKKTLSVSQLELDQSG 180
QY 181 TWTCTVLQNGKVEFKIDIVLAFOKAS 208
DB 181 TWTCTVLQNGKVEFKIDIVLAFOKAS 208

```

```

RESULT 8
US-10-820-403-25
; Sequence 25, Application US/10820403
; Publication No. US20040229287A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Taka-Aki
; TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
; TITLE OF INVENTION: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: 65823/JPW/PT

```

CURRENT APPLICATION NUMBER: US/10/820,403  
CURRENT FILING DATE: 2004-04-08  
PRIOR APPLICATION NUMBER: US/10/092,138  
PRIOR FILING DATE: 2002-03-06  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 25  
LENGTH: 458  
TYPE: PRT  
ORGANISM: human  
US-10-820-403-25

Query Match 39.8%; Score 1062; DB 16; Length 458;  
Best Local Similarity 99.5%; Pred. No. 9.5e-51;  
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLLVQLALPAATQGRVVLGKGGDTVELTCTASQKSIQFHKNSNQIK 60  
DB 1 MNRGVPFRHLVLLVQLALPAATQGRVVLGKGGDTVELTCTASQKSIQFHKNSNQIK 60  
QY 61 ILGNQGSFLTKGSPSKLNDRAISRSLMDQGNFPLIINKLIKIDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLTKGSPSKLNDRAISRSLMDQGNFPLIINKLIKIDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLQGSQSLTILTESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
DB 121 LVFGLTANSDTHLQGSQSLTILTESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
QY 181 TWTCTVLQNKQKVEFKIDIVLAFQKAS 208  
DB 181 TWTCTVLQNKQKVEFKIDIVLAFQKAS 208

## RESULT 9

US-09-939-537-29  
Sequence 29, Application US/09939537  
Publication No. US20030138410A1  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
Banapour, Babak  
Romeo, Charles  
Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INJECTED  
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbling LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,537  
FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbling, Karen L.  
REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 398 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 29:  
US-09-939-537-29

Query Match 39.7%; Score 1059; DB 10; Length 398;  
Best Local Similarity 99.5%; Pred. No. 1.2e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLLVQLALPAATQGRVVLGKGGDTVELTCTASQKSIQFHKNSNQIK 60  
DB 1 MNRGVPFRHLVLLVQLALPAATQGRVVLGKGGDTVELTCTASQKSIQFHKNSNQIK 60  
QY 61 ILGNQGSFLTKGSPSKLNDRAISRSLMDQGNFPLIINKLIKIDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLTKGSPSKLNDRAISRSLMDQGNFPLIINKLIKIDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLQGSQSLTILTESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
DB 121 LVFGLTANSDTHLQGSQSLTILTESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
QY 181 TWTCTVLQNKQKVEFKIDIVLAFQKAS 208  
DB 181 TWTCTVLQNKQKVEFKIDIVLAFQKAS 208

## RESULT 10

US-10-157-408-1  
Sequence 1, Application US/10157408  
Publication No. US20030104535A1  
GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
Gregory, Timothy J.  
TITLE OF INVENTION: Adheson Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/157,408  
FILING DATE: 28-May-2002  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,918  
FILING DATE: 1-JUN-1995  
APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
APPLICATION NUMBER: 07/104329

FILING DATE: 02-OCT-1987  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kubinec, Jeffrey S.  
 REGISTRATION NUMBER: 36,575  
 REFERENCE/DOCKET NUMBER: P0444P1C3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415/225-8228  
 TELEFAX: 415/952-9881  
 TELEX: 910/371-7168  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 402 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-10-157-408-1

Query Match 39.7%; Score 1059; DB 14; Length 402;  
 Best Local Similarity 99.5%; Pred. No. 1.2e-50;  
 Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLTQALLPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60  
 DB 1 MNRGVPFRHLVLTQALLPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60  
 QY 61 ILNGQSFLLTKGSKINDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEQKEEYOL 120  
 DB 61 ILNGQSFLLTKGSKINDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEQKEEYOL 120  
 QY 121 LVFGLTANSDTHLLQGSLLTLETSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDTHLLQGSLLTLETSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWTCTVQLNQKQKVEFKIDIVLAFOKAS 208  
 DB 181 TWTCTVQLNQKQKVEFKIDIVLAFOKAS 208

RESULT 11  
 US-10-097-044A-1  
 ; Sequence 1, Application US/10097044A  
 ; Publication No. US20030143220A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Capon, Daniel J.  
 ; TITLE OF INVENTION: Adhesion Variants  
 ; NUMBER OF SEQUENCES: 25  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genentech, Inc.  
 ; STREET: 460 Point San Bruno Blvd  
 ; CITY: South San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94080  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: patin (Genentech)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/097,044A  
 ; FILING DATE: 28-May-2002  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/457,918  
 ; FILING DATE: 1-JUN-1995  
 ; APPLICATION NUMBER: 08/236311  
 ; FILING DATE: 02-MAY-1994  
 ; APPLICATION NUMBER: 07/936190  
 ; FILING DATE: 26-AUG-1992  
 ; APPLICATION NUMBER: 07/842777  
 ; FILING DATE: 18-FEB-1992  
 ; APPLICATION NUMBER: 07/250785

FILING DATE: 28-SEP-1988  
 APPLICATION NUMBER: 07/104329  
 FILING DATE: 02-OCT-1987  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kubinec, Jeffrey S.  
 REGISTRATION NUMBER: 36,575  
 REFERENCE/DOCKET NUMBER: P0444P1C3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415/225-8228  
 TELEFAX: 415/952-9881  
 TELEX: 910/371-7168  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 402 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-10-097-044A-1

Query Match 39.7%; Score 1059; DB 14; Length 402;  
 Best Local Similarity 99.5%; Pred. No. 1.2e-50;  
 Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLTQALLPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60  
 DB 1 MNRGVPFRHLVLTQALLPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60  
 QY 61 ILNGQSFLLTKGSKINDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEQKEEYOL 120  
 DB 61 ILNGQSFLLTKGSKINDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEQKEEYOL 120  
 QY 121 LVFGLTANSDTHLLQGSLLTLETSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDTHLLQGSLLTLETSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWTCTVQLNQKQKVEFKIDIVLAFOKAS 208  
 DB 181 TWTCTVQLNQKQKVEFKIDIVLAFOKAS 208

RESULT 12  
 US-10-769-247-1  
 ; Sequence 1, Application US/10769247  
 ; Publication No. US20040197809A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Capon, Daniel J.  
 ; TITLE OF INVENTION: Adhesion Variants  
 ; NUMBER OF SEQUENCES: 25  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genentech, Inc.  
 ; STREET: 460 Point San Bruno Blvd  
 ; CITY: South San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94080  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: patin (Genentech)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/769,247  
 ; FILING DATE: 30-Jan-2004  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/457,918  
 ; FILING DATE: 1-JUN-1995  
 ; APPLICATION NUMBER: 08/236311  
 ; FILING DATE: 02-MAY-1994  
 ; APPLICATION NUMBER: 07/936190  
 ; FILING DATE: 26-AUG-1992  
 ; APPLICATION NUMBER: 07/842777

```

; FILING DATE: 18-FEB-1992
; APPLICATION NUMBER: 07/250785
; FILING DATE: 28-SEP-1988
; APPLICATION NUMBER: 07/104329
; FILING DATE: 02-OCT-1987
; ATTORNEY/AGENT INFORMATION:
;   NAME: Kubinec, Jeffrey S.
;   REGISTRATION NUMBER: 36,575
;   REFERENCE/DOCKET NUMBER: P0444P1C3
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415/225-8228
;   TELEFAX: 415/952-9881
;   TELEEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 402 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-769-247-1

```

```

Query Match      39.7%; Score 1059; DB 16; Length 402;
Best Local Similarity 99.5%; Pred. No. 1,2e-50;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MNRGVPFRHLHLVQLALLPAATGGKVVLGKKGDTVELTCTASQKKSIOFHMKNNSQIK 60
DB 1 MNRGVPFRHLHLVQLALLPAATGGKVVLGKKGDTVELTCTASQKKSIOFHMKNNSQIK 60
QY 61 ILNGSGFLTKGSPSKNDRAADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120
DB 61 ILNGSGFLTKGSPSKNDRAADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTANSDTHLLQGOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLISVQLELQDSG 180
DB 121 LVFGLTANSDTHLLQGOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLISVQLELQDSG 180
QY 181 TWCTVVLQNGKVEFKIDIVLAFQKAS 208
DB 181 TWCTVVLQNGKVEFKIDIVLAFQKAS 208

```

```

RESULT 13
US-09-891-119A-9
; Sequence 9, Application US/09891119A
; Publication No. US20040013683A1
; GENERAL INFORMATION:
;   APPLICANT: Maddon, Paul J.
;   TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4
;   FILE REFERENCE: 24577-CY-B
;   CURRENT APPLICATION NUMBER: US/09/891,119A
;   CURRENT FILING DATE: 2001-06-25
;   NUMBER OF SEQ ID NOS: 22
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
;   LENGTH: 457
;   TYPE: PRT
;   ORGANISM: human
US-09-891-119A-9

```

```

Query Match      39.7%; Score 1059; DB 11; Length 457;
Best Local Similarity 99.5%; Pred. No. 1,4e-50;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MNRGVPFRHLHLVQLALLPAATGGKVVLGKKGDTVELTCTASQKKSIOFHMKNNSQIK 60
DB 1 MNRGVPFRHLHLVQLALLPAATGGKVVLGKKGDTVELTCTASQKKSIOFHMKNNSQIK 60
QY 61 ILNGSGFLTKGSPSKNDRAADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120
DB 61 ILNGSGFLTKGSPSKNDRAADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTANSDTHLLQGOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLISVQLELQDSG 180

```

```

DB 121 LVFGLTANSDTHLLQGOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLISVQLELQDSG 180
QY 181 TWCTVVLQNGKVEFKIDIVLAFQKAS 208
DB 181 TWCTVVLQNGKVEFKIDIVLAFQKAS 208

```

```

RESULT 14
US-09-939-537-5
; Sequence 5, Application US/09939537
; Publication No. US20030138410A1
; GENERAL INFORMATION:
;   APPLICANT: Seed, Brian
;             Banapour, Babak
;             Romeo, Charles
;             Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
;                   CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Clark & Elbing LLP
;   STREET: 176 Federal Street
;   CITY: Boston
;   STATE: MA
;   COUNTRY: USA
;   ZIP: 02110
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/939,537
;   FILING DATE: 24-Aug-2001
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/284,391
;   FILING DATE: 02-AUG-1994
;   APPLICATION NUMBER: 08/195,395
;   FILING DATE: 14-FEB-1994
;   APPLICATION NUMBER: 07/847,566
;   FILING DATE: 06-MAR-1992
;   APPLICATION NUMBER: 07/665,961
;   FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
;   NAME: Elbing, Karen L.
;   REGISTRATION NUMBER: 35,238
;   REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617-428-0200
;   TELEFAX: 617-428-7045
;   TELEEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 462 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-939-537-5

```

```

Query Match      39.7%; Score 1059; DB 10; Length 462;
Best Local Similarity 99.5%; Pred. No. 1,4e-50;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MNRGVPFRHLHLVQLALLPAATGGKVVLGKKGDTVELTCTASQKKSIOFHMKNNSQIK 60
DB 1 MNRGVPFRHLHLVQLALLPAATGGKVVLGKKGDTVELTCTASQKKSIOFHMKNNSQIK 60
QY 61 ILNGSGFLTKGSPSKNDRAADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120

```

Db 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIRKLIKIEDSDTYICEVEDQKEVQL 120  
Qy 121 LVFGLTANSPDTHLQGSGLTLTLTLESPGSSPSVQCRSPRGKNIQGGKTLTSSVQLLEODSG 180  
Db 121 LVFGLTANSPDTHLQGSGLTLTLTLESPGSSPSVQCRSPRGKNIQGGKTLTSSVQLLEODSG 180  
Qy 181 TWTCTVLQONOKKVEFKIDIVLAFQKAS 208  
Db 181 TWTCTVLQONOKKVEFKIDIVLAFQKAS 208

## RESULT 15

US-09-243-008-5  
; Sequence 5, Application US/09243008  
; Publication No. US20040005334A1  
; GENERAL INFORMATION:  
; APPLICANT: Seed, Brian et al.  
; TITLE OF INVENTION: Redirection of Cellular Immunity by  
; Receptor Chimeras  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
; SOFTWARE: Wordperfect (Version 5.0)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/243,008  
; FILING DATE: 02-Feb-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/394,176  
; FILING DATE: SEPTEMBER 11, 1995  
; APPLICATION NUMBER: 08/203,866  
; FILING DATE: February 28, 1994  
; APPLICATION NUMBER: 07/847,566  
; FILING DATE: March 6, 1992  
; APPLICATION NUMBER: 07/665,961  
; FILING DATE: March 7, 1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Karen F. Iech, Ph.D.  
; REGISTRATION NUMBER: 35,238  
; REFERENCE/DOCKET NUMBER: 00786/270001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 462 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
; US-09-243-008-5

Query Match 39.7%; Score 1059; DB 11; Length 462;  
Best Local Similarity 99.5%; Pred. No. 1.4e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MNRGVPFRHLVQLALPAPATQKRVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60  
Db 1 MNRGVPFRHLVQLALPAPATQKRVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60  
Qy 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIRKLIKIEDSDTYICEVEDQKEVQL 120  
Db 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIRKLIKIEDSDTYICEVEDQKEVQL 120

Qy 121 LVFGLTANSPDTHLQGSGLTLTLTLESPGSSPSVQCRSPRGKNIQGGKTLTSSVQLLEODSG 180  
Db 121 LVFGLTANSPDTHLQGSGLTLTLTLESPGSSPSVQCRSPRGKNIQGGKTLTSSVQLLEODSG 180  
Qy 181 TWTCTVLQONOKKVEFKIDIVLAFQKAS 208  
Db 181 TWTCTVLQONOKKVEFKIDIVLAFQKAS 208

Search completed: September 20, 2005, 16:27:40  
Job time : 487 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 13:43:29 ; Search time 34 Seconds

(without alignments)  
1437.592 Million cell updates/sec

Title: US-09-936-702-3

Perfect score: 2668

Sequence: 1 MNRGVPRRLLLVLTALLP.....GTRLEIKLVPRGSGHHHHH 508

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1064	39.9	458	1	RHMT4
2	916	34.3	432	1	RMCZT4
3	831	31.1	432	1	RMOT4
4	750	28.1	268	2	A56446
5	668	25.0	249	2	S41374
6	618.5	23.2	459	2	A46254
7	616	23.1	233	2	JC5322
8	596	22.3	432	2	S30193
9	508.5	19.1	457	2	A27449
10	507	19.0	128	2	S40343
11	498	18.7	107	2	S34005
12	498	18.7	117	2	S40362
13	498	18.7	457	1	RMMS74
14	485.5	18.2	116	2	B26554
15	485	18.2	123	2	S40378
16	481.5	18.0	131	2	S40358
17	480.5	18.0	109	1	K3HUPM
18	475.5	17.8	215	2	JE0244
19	466	17.5	111	2	S23628
20	460	17.2	108	2	G44151
21	458	17.2	128	2	A56701
22	456.5	17.1	114	2	S54905
23	456	17.1	128	2	S40345
24	455	17.1	114	2	S46375
25	455	17.1	129	2	S29627
26	454	17.0	115	1	K3HUCT
27	453	17.0	98	2	I30608
28	453	17.0	115	2	A30553
29	453	17.0	144	2	PL0106

30	451	16.9	95	2	PH0868	Ig kappa chain V r
31	451	16.9	215	2	JE0243	Ig kappa chain NIG
32	449	16.8	110	2	S40326	Ig kappa chain V-J
33	448.5	16.8	109	2	S30601	Ig kappa chain V-I
34	448	16.8	128	2	S40379	Ig kappa chain V-J
35	445.5	16.7	627	2	S14683	Ig mu chain precu
36	443.5	16.6	108	2	C30608	Ig kappa chain V-I
37	441.5	16.5	129	2	S46369	Ig light chain var
38	441.5	16.5	134	2	S38643	Ig kappa chain V r
39	440.5	16.5	107	2	PH0965	Ig kappa chain V r
40	440.5	16.5	109	2	PH0963	Ig kappa chain V r
41	438.5	16.4	109	2	A30608	Ig kappa chain V-I
42	438.5	16.4	109	2	F30607	Ig kappa chain V-I
43	438	16.4	107	2	S57444	Ig kappa chain V-J
44	437.5	16.4	109	2	B30601	Ig kappa chain V-I
45	436.5	16.4	109	2	G30601	Ig kappa chain V-I

#### ALIGNMENTS

##### RESULT 1

RHMT4  
T-cell surface glycoprotein CD4 precursor [validated] - human  
N:Alternate names: T-cell surface antigen T4/Len 3  
C:Species: Homo sapiens (man)  
C>Date: 28-May-1986 #sequence, revision 31-Dec-1988 #text change 09-Jul-2004  
C:Accession: A90872; A32722; A34194; A53287; I54176; I54297; A02109; A30039  
R:Maddon, P.J.; Liteman, D.R.; Godfrey, M.; Maddon, D.E.; Cheese, L.; Axel, R.  
Cell 42, 93-104, 1985  
A:Title: The isolation and nucleotide sequence of a cDNA encoding the T cell surface protein  
A:Reference number: A90872; PMID:85254948; PMID:2990730  
A:Accession: A90872  
A:Molecule type: mRNA  
A:Residues: 1-25, 'N', 27-458 <MAD>  
A:Cross-references: UNIPROT:P01730  
A:Experimental source: clone PT4B  
R:Liteman, D.R.; Maddon, P.J.; Axel, R.  
Cell 55, 541, 1988  
A:Title: Corrected CD4 sequence.  
A:Reference number: A90807; PMID:89028665; PMID:3263213  
A:Contents: annotation; revision to residue 26  
R:Camerini, D.; Seed, B.  
Cell 60, 747-754, 1990  
A:Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi  
A:Reference number: A32722; PMID:90182664; PMID:2107024  
A:Accession: A32722  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 26-426, 428-458 <CAM>  
R:Carri, S.A.; Henling, M.B.; Folema-Wasserman, G.; Sweet, R.W.; Anumula, K.; Barr, J.R.;  
J. Biol. Chem. 264, 21286-21295, 1989  
A:Title: Protein and carbohydrate structural analysis of a recombinant soluble CD4 recept  
A:Reference number: A34194; PMID:90078232; PMID:2592374  
A:Contents: disulfide bonds; carbohydrate-binding sites  
A:Accession: A34194  
A:Molecule type: protein  
A:Residues: 26-354 <CAR>  
R:Lederman, S.; DeMartino, J.A.; Daugherty, B.L.; Foeldvari, I.; Yellin, M.J.; Cleary, A.  
Mol. Immunol. 28, 1171-1181, 1991  
A:Title: A single amino acid substitution in a common African allele of the CD4 molecule  
A:Reference number: A53287; PMID:92072595; PMID:1961196  
A:Accession: A53287  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 250-264, 'W', 266-280 <LED>  
A>Note: sequence extracted from NCBI backbone (NCBIP.68249)  
R:Edwards, M.C.; Gibbs, R.A.  
Genomics 14, 590-597, 1992  
A:Title: A human dimorphism resulting from loss of an Alu.  
A:Reference number: I54176; PMID:93052367; PMID:1330888  
A:Accession: I54176  
A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-72 <RES>  
 A:Cross-references: GB:U47924; GB:M86525; GB:U72506; NID:g1633547; PIDN:AAB51309.1; PID:R.Hodge, T.W.; Sasso, D.R.; McConugal, J.S.  
 Hum. Immunol. 30, 99-104, 1991  
 A:Title: Humans with OK4-epitope deficiency have a single nucleotide base change in the A:Reference number: I54297; MUID:91216786; PMID:11708753  
 A:Accession: I54297  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-264 'W', 266-458 <RES>  
 A:Cross-references: GB:M35160; NID:g179143; PIDN:AAA16069.1; PID:g179144  
 A:Comment: Macrophage tropic strains of HIV-1 bind to a complex of chemokine (C-C) receptor C:Genetics:  
 A:Gene: GDB:CD4  
 A:Cross-references: GDB:119767; OMIM:186940  
 A:Map position: 12pter-12p12  
 A:Features: 16/3  
 C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C:Keywords: AIDS; duplication; glycoprotein; T-cell; transmembrane protein  
 F:1-25/Domain: signal sequence #status predicted <SIG>  
 F:26-458/Product: T-cell surface glycoprotein CD4 #status experimental <MAT>  
 F:34-111/Domain: immunoglobulin homology <IM1>  
 F:136-186/Domain: immunoglobulin homology #status atypical <IM2>  
 F:216-299/Domain: immunoglobulin homology <IM3>  
 F:321-372/Domain: immunoglobulin homology <IM4>  
 F:397-420/Domain: transmembrane #status predicted <TM>  
 F:421-458/Domain: intracellular #status predicted <INT>  
 F:41-109,155-184,328-370/Disulfide bonds: #status experimental  
 F:266,325/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 39.9%; Score 1064; DB 1; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 1e-48;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALLPAATGKRVYLGKKSDTYELTCTASQKSIQPHMKNNOIK 60  
 DB 1 MNRGVPFRHLLVQLALLPAATGKRVYLGKKSDTYELTCTASQKSIQPHMKNNOIK 60  
 QY 61 ILNQGSLFKGSPKLNDRADSRSLMDQGNPFIILNKLEISDITYICEVEDQKEVQL 120  
 DB 61 ILNQGSLFKGSPKLNDRADSRSLMDQGNPFIILNKLEISDITYICEVEDQKEVQL 120  
 QY 121 LVFGLTANSDTHLLQGSLTLTLESPGSSPVQCRSPRKNIOGKTLTSLVSQLELDGSG 180  
 DB 121 LVFGLTANSDTHLLQGSLTLTLESPGSSPVQCRSPRKNIOGKTLTSLVSQLELDGSG 180  
 QY 181 TWTCTVLQONKXVEFKIDIVVLAFOKAS 208  
 DB 181 TWTCTVLQONKXVEFKIDIVVLAFOKAS 208

## RESULT 2

RWC274  
 T-cell surface glycoprotein CD4 - chimpanzee  
 N:Alternate names: T-cell surface antigen T4/Lew 3  
 C:Species: Pan troglodytes (chimpanzee)  
 C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C:Accession: B32722; A46534  
 R:Cametini, D.; Seed, B.  
 Cell 60, 747-754, 1990  
 A:Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi  
 A:Reference number: A32722; MUID:90182664; PMID:2107024  
 A:Accession: B32722  
 A:Molecule type: mRNA  
 A:Residues: 1-432 <CAM>  
 A:Cross-references: UNIPROT:P16004; GB:M11135  
 R:Fomsgaard, A.; Hirsch, V.M.; Johnson, P.R.  
 Eur. J. Immunol. 22, 2973-2981, 1992  
 A:Title: Cloning and sequences of primate CD4 molecules: diversity of the cellular recep  
 A:Reference number: A46534; MUID:93049640; PMID:11455921  
 A:Accession: A46534  
 A:Status: not compared with conceptual translation

A:Molecule type: mRNA  
 A:Residues: 3-399 <FOM>  
 A:Note: sequence extracted from NCBI backbone (NCBIP:118332)  
 C:Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells ct  
 C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; T-cell; transmembrane protein  
 F:1-432/Product: T-cell surface glycoprotein CD4 #status predicted <MAT>  
 F:1-371/Domain: extracellular #status predicted <EXT>  
 F:9-86/Domain: immunoglobulin homology <IM1>  
 F:111-161/Domain: immunoglobulin homology #status atypical <IM2>  
 F:191-274/Domain: immunoglobulin homology <IM3>  
 F:286-347/Domain: immunoglobulin homology <IM4>  
 F:312-395/Domain: transmembrane #status predicted <TM>  
 F:396-432/Domain: intracellular #status predicted <INT>  
 F:16-84,130-159,303-345/Disulfide bonds: #status predicted  
 F:271,300/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 34.3%; Score 916; DB 1; Length 432;  
 Best Local Similarity 97.8%; Pred. No. 4.9e-41;  
 Matches 179; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 26 KKVVLGKKGDYELTCTASQKSIQPHMKNNOIKILNQGSLFKGSPKLNDRADSR 85  
 DB 1 KKVVLGKKGDYELTCTASQKSIQPHMKNNOIKILNQGSLFKGSPKLNDRADSR 85  
 QY 86 LMDQGNPFIILNKLEISDITYICEVEDQKEVQLVFGLTANSDTHLLQGSLTLTLES 145  
 DB 61 LMDQGNPFIILNKLEISDITYICEVEDQKEVQLVFGLTANSDTHLLQGSLTLTLES 120  
 QY 146 PPGSSPVQCRSPRKNIOGKTLTSLVSQLELDGSGTWTCTVLQONKXVEFKIDIVVLAQ 205  
 DB 121 PPGSSPVQCRSPRKNIOGKTLTSLVSQLELDGSGTWTCTVLQONKXVEFKIDIVVLAQ 180  
 QY 206 KAS 208  
 DB 181 KAS 183

## RESULT 3

RWC274  
 T-cell surface glycoprotein CD4 - rhesus macaque  
 N:Alternate names: T-cell surface antigen T4/Lew 3  
 C:Species: Macaca mulatta (rhesus macaque)  
 C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Jul-1999  
 C:Accession: C32722  
 R:Cametini, D.; Seed, B.  
 Cell 60, 747-754, 1990  
 A:Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi  
 A:Reference number: A32722; MUID:90182664; PMID:2107024  
 A:Accession: C32722  
 A:Molecule type: mRNA  
 A:Residues: 1-432 <CAM>  
 A:Cross-references: GB:M11134  
 C:Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells ct  
 C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; T-cell; transmembrane protein  
 F:1-432/Product: T-cell surface glycoprotein CD4 #status predicted <MAT>  
 F:1-371/Domain: extracellular #status predicted <EXT>  
 F:9-86/Domain: immunoglobulin homology <IM1>  
 F:111-161/Domain: immunoglobulin homology #status atypical <IM2>  
 F:180-293/Domain: immunoglobulin homology <IM3>  
 F:286-347/Domain: immunoglobulin homology <IM4>  
 F:312-395/Domain: transmembrane #status predicted <TM>  
 F:396-432/Domain: intracellular #status predicted <INT>  
 F:16-84,130-159,303-345/Disulfide bonds: #status predicted  
 F:271,300/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 31.1%; Score 831; DB 1; Length 432;  
 Best Local Similarity 88.0%; Pred. No. 1.3e-36;  
 Matches 161; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY 26 KKVVLGKKGDYELTCTASQKSIQPHMKNNOIKILNQGSLFKGSPKLNDRADSR 85  
 DB 1 KKVVLGKKGDYELTCTASQKSIQPHMKNNOIKILNQGSLFKGSPKLNDRADSR 85



Db 1 KKVVLGKKGDVTELTCTASOKKNTQFHWKNSNOIKIIGQLFLTKGPSKLSDRADSRKS 60  
 QY 86 LMDGNFPLIITKNKIEDSDTYICEVEDQKEVOLVGLFANSDDTHLQGSLLTLTES 145  
 Db 61 LMDGSCSMIITKNKIEDSDTYICEVENKKEVELVGLFANSDDTHLQGSLLTLTES 120  
 QY 146 PGGSSPVOCSPRGKNIQGGKTLVSQLELQDSGTWCTVLQNKVFEKIDIVLAFQ 205  
 Db 121 PGGSSPVOCSPRGKNIQGGKTLVSQLELQDSGTWCTVLQNKVFEKIDIVLAFQ 180  
 QY 206 KAS 208  
 Db 181 KAS 183

## RESULT 4

A56446  
 Ig heavy chain V region (3H-3H scFv) - mouse (strain BALB/C)  
 C/Species: Mus musculus (house mouse)  
 C/Date: 19-Jan-1996 #sequence\_revision 19-Jan-1996 #text\_change 16-Aug-1996  
 C/Accession: A56446  
 R/Tang, P.M.; Foltz, L.A.; Mahoney, W.C.; Schueler, P.A.  
 J. Biol. Chem. 270, 7829-7835, 1995  
 A/Title: A high affinity digoxin-binding protein displayed on M13 is functionally identical  
 A/Reference number: A56446; MUID:95229583; PMID:7713873  
 A/Accession: A56446  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-268 <TAN>  
 A/Cross-references: GB:U20617  
 C/Keywords: heterotrimer; immunoglobulin

Query March 28.1%; Score 750; DB 2; Length 268;  
 Best Local Similarity 55.6%; Pred. No. 1.4e-32;  
 Matches 150; Conservative 39; Mismatches 57; Indels 24; Gaps 7;  
 QY 243 SQVQLLESGAEVKKPGSSGVKSCAKSG---DTEIRYFTVWROAPQGLEMMGRITITL 298  
 Db 2 AQVQLQSGAEVKKPGSSGVKSCAKSG---DTEIRYFTVWROAPQGLEMMGRITITL 298  
 QY 299 DVAAHAPHLQGRVITTDKSTSTVYLELRNRSDDTAAYFCAGYEGEADGEYDNNQFL 358  
 Db 58 GITKCDKPFQKATITADTSNTAYLQLSLTSEDTAAYVYCASYY---LTRYEN---- 108  
 QY 359 KHWQGTLYVTSSGGSGSGGSGGSELELTOSPATLSVPERATLSGRABESVSD 418  
 Db 109 -YMQQGTLYVTSSGGSGSGGSGGSDIELTQSPAIMSASLDEKVTMGRASSVNF- 166  
 QY 419 LAWYQKPGQAPRLIYGASTRATGVPARFSGSGGAEFTLTISLQSEDPFAVYCCOYN 478  
 Db 167 IYWIQKSDASPKLWVYTHSLPRGVPARFSGSGSGNSYSLTISMEGEDATYYCCOFT 226  
 QY 479 NMPRYTFQGTTRLEIKLVPRSGHHHHH 508  
 Db 227 SSP--FTFGSGTKLEIK---RSAHHHHH 251

## RESULT 5

S41374  
 single chain Fv antibody - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 06-Jan-1995  
 C/Accession: S41374  
 R/Atsuenko, O.; Weiler, E.W.; Muentz, K.; Conrad, U.  
 submitted to the EMBL Data Library, January 1994  
 A/Description: Construction and functional characterization of a single chain Fv antibody  
 A/Reference number: S41374  
 A/Accession: S41374  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-249 <ART>  
 A/Cross-references: EMBL:Z29480

Query March 25.0%; Score 668; DB 2; Length 249;  
 Best Local Similarity 52.9%; Pred. No. 2.5e-28;  
 Matches 136; Conservative 40; Mismatches 65; Indels 16; Gaps 4;  
 QY 244 QVQLLESGAEVKKPGSSGVKSCAKSGDTPRISFTVWROAPQGLEMMGRITITLVAHY 303  
 Db 1 QVQLQSGAEVKKPGSSGVKSCAKSGFNDQYIHAWKQPERGGLMIAIAPASGVNKY 60  
 QY 304 APHLQGRVITTDKSTSTVYLELRNRSDDTAAYFCAGYEGEADGEYDNNQFLKHWQ 363  
 Db 61 VPRPDATITADTSNTAYLQLSLTSEDTAAYVYCA-----RDPLTYSGLY---WQ 111  
 QY 364 GTLVTVSGGSGSGGSGGSGGSELELTOSPATLSVPERATLSGRABESV-----SSD 418  
 Db 112 GSTVTVSSRGSGSGGSGGSGGSDIELTQSPVAVIPGSSVLSICRSSSLYSQSDSY 171  
 QY 419 LAWYQKPGQAPRLIYGASTRATGVPARFSGSGGAEFTLTISLQSEDPFAVYCCOYN 478  
 Db 172 LFWFLQRPQSPQLLIYRMSNLASGVDPDRSGSGSGFTLRISVBAEDVGYVYCCOHR 231  
 QY 479 NMPRYTFQGTTRLEIK 495  
 Db 232 EYP--LTFGAGTYLEIK 246

## RESULT 6

A46254  
 CD4 precursor - rabbit  
 C/Species: Oryctolagus cuniculus (domestic rabbit)  
 C/Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
 C/Accession: A46254  
 R/Hague, B.F.; Sawadkoso, S.; Brown, T.J.; Lee, K.; Recker, D.P.; Kindt, T.J.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 7963-7967, 1992  
 A/Title: CD4 and its role in infection of rabbit cell lines by human immunodeficiency virus  
 A/Reference number: A46254; MUID:92390370; PMID:1518821  
 A/Accession: A46254  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-459 <HAG>  
 A/Cross-references: UNIPROT:P46630; GB:M92840; NID:g164871; PIND:AAA31198.1; PID:g164872  
 A/Note: sequence extracted from NCBI backbone (NCBI:112732, NCBI:P:112733)  
 C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 F:322-372/Domain: immunoglobulin homology <IMM>

Query March 23.2%; Score 618.5; DB 2; Length 459;  
 Best Local Similarity 59.6%; Pred. No. 1.7e-25;  
 Matches 127; Conservative 34; Mismatches 47; Indels 5; Gaps 2;  
 QY 1 MNRGVPFRHLVLYLALPAATQGGKVVYGGKDTVELTCTASOKKSIQFHWKNSNOIK 60  
 Db 1 MNRRIYFQCLLVLPALPAAITWGTIVYRGKAGALVELPCQSSQKNSVFMKHAQVK 60  
 QY 61 ILNQG-----SPLTKGPSKLNDRADSRRLMDGNFPLIITKNKIEDSDTYICEVEDQKE 116  
 Db 61 ILNQGSSSSFWLKGNSPLSNRVESKQNMWDGSPFLVTKDRLMDSGTYICEVEDQKM 120  
 QY 117 EVQLVFPGLTANSDTHLQGSITLTLESPPGSSPSVOCSPRGKNIQGGKTLVSQLEL 176  
 Db 121 EVELVFRLLTANSDTHLQGSITLTLEGPSSPSVOCSPRGKNIQGGKTLVSQLEL 180  
 QY 177 QDSGTWCTV-LQNKVFEKIDIVLAFQKAS 208  
 Db 181 QDSGTWCTV-LQNKVFEKIDIVLAFQKAS 213

## RESULT 7

JC5322  
 p53 specific single-chain antibody Pab421 - human  
 C/Species: Homo sapiens (man)  
 C/Date: 15-May-1997 #sequence\_revision 15-May-1997 #text\_change 18-Jul-1997  
 C/Accession: JC5322  
 R/Jannot, C.B.; Hynes, N.E.  
 Biochem. Biophys. Res. Commun. 230, 242-246, 1997

A>Title: Characterization of acFv-421, a single-chain antibody targeted to p53.  
 A/Reference number: JCS322; MUID:97168950; PMID:9016757  
 A/Accession: JCS322  
 A/Molecule type: mRNA  
 A/Residues: 1-233 <JAN>  
 A/Experimental source: hydridicoma cell  
 C/Comment: This protein specifically binds the tumor suppressor protein p53. It restores

Query Match 23.1%; Score 616; DB 2; Length 233;  
 Best Local Similarity 52.6%; Pred. No. 1.2e-25;  
 Matches 132; Conservative 32; Mismatches 63; Indels 24; Gaps 4;

QY 249 ESGAEVKKPPGSSVSVSCASDPTIRKFTVWRAPQSGLEMMKRITITLDVAHYAHLQ 308  
 DB 2 ESGAEVRSQASVYLSCTTSQFNINDYMMWKKRPPQGLWICRIPENDGDMTRSSG 61  
 QY 309 GRVITTDKSTSYLYELRNRSDDTVYFCAGYEEGADGEVDNNGLKHMGGTLYT 368  
 DB 62 VKATMTDTSNTYLYQLSSLTSEDAVYIC-----NAGMDYMGQGTITVT 106  
 QY 369 VTSGGGSGGGGGSGGSELELTQSPATLSVSPERATLSCRASESYS---SDLAWYQQ 424  
 DB 107 VSSGGSGGSGRAGSGGSDIELTQSPASLAVSLGQRAITISCRASKSVSTSGSYMHNNQ 166  
 QY 425 KPGQAPRLIYGASTRATGVPARFSGSGGAEFTLTSSLOSEPFAVYCCQINNMPRY 484  
 DB 167 KPGQAPRLIYLVNLSGVPARFSGSGGSDFTLTINHPVEEDAAATYCCQHIR---EL 222  
 QY 485 TFGQ-GTRLEI 494  
 DB 223 TRSEGTRLEI 233

## RESULT 8

S30193  
 T-cell surface glycoprotein CD4 - dog  
 C/Species: Canis lupus familiaris (dog)  
 C/Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text\_change 21-Jul-2000  
 C/Accession: S30193  
 R/Mide, K.F.; Conner, G.E.; Mintz, D.H.; Alejandro, R.  
 Biochim. Biophys. Acta 1172, 315-318, 1993  
 A>Title: Primary structure of the canine CD4 antigen.  
 A/Reference number: S30193; MUID:93192324; PMID:7916632  
 A/Accession: S30193  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-432 <MIL>  
 A/Cross-references: EMBL:X68565; NID:9288652; PIDD:CA837664.1; PIDD:g4467377  
 C/Superfamily: T-cell surface glycoprotein CD4; Immunoglobulin homology  
 C/Keywords: glycoprotein  
 F/202-311/Domain: immunoglobulin homology <IMM>

Query Match 22.3%; Score 596; DB 2; Length 432;  
 Best Local Similarity 56.3%; Pred. No. 2.4e-24;  
 Matches 116; Conservative 38; Mismatches 42; Indels 10; Gaps 2;

QY 12 LVLTQALLPAPATQKRVVGGKGDVETLTASQKSIQPHMKNSNOIKILGNGSFYTK 71  
 DB 1 LMTQVLWLPATVPREVVLGKADAVELPCQTSQKNIHFMMRDSMWQILGNGSFWTV 60  
 QY 72 GPEGLNDRADSRRLMQNGNPLIILKULKIEDSTTYICEVEDQKEEYQVLVGLTA--- 127  
 DB 61 GSSRLKRRVSKKMLMWDGSPFLVYIKDLVADSGIYCDT-DKQOEVELLVFNLTATWDS 119  
 QY 128 -----NSDTLHLOGSLTLTLESPGSSPSVQCRSPGKNIQGGKTLVSQLELDGSGTW 182  
 DB 120 GSSSGSSNNIRLLQOQQLTLTENPSSGSSPSVQMGKGNKSGHGGNLSLWPELDGGIN 179  
 QY 183 TCTVLQNGKVEFKIDIVVLAFOKAS 208  
 DB 180 TCTISGQKVEFNIIVVLAFOKAS 205

RESULT 9  
 A27449  
 T-cell surface glycoprotein CD4 precursor - rat

N/Alternate names: W3/25 antigen  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 21-May-1988 #sequence revision 21-May-1988 #text\_change 09-Jul-2004  
 C/Accession: A27449; A35433  
 R/Clark, S.J.; Jelfer, W.A.; Barclay, A.N.; Gagnon, J.; Williams, A.F.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 1649-1653, 1987  
 A>Title: Peptide and nucleotide sequences of rat CD4 (W3/25) antigen: evidence for derive  
 A/Reference number: A27449; MUID:8715535; PMID:3104900  
 A/Accession: A27449  
 A/Molecule type: mRNA  
 A/Residues: 1-457 <CIA>  
 A/Cross-references: UNIPROT:P05540; GB:M15768; NID:9203387; PIDD:AAA0901.1; PIDD:9203388  
 R/Davis, S.J.; Ward, H.A.; Puklavec, M.J.; Willis, A.C.; Williams, A.F.; Barclay, A.N.  
 J. Biol. Chem. 265, 10410-10418, 1990  
 A>Title: High level expression in Chinese hamster ovary cells of soluble forms of CD4 T  
 A/Content: annotation  
 A/Reference number: A35433; MUID:90285164; PMID:2113054  
 C/Superfamily: T-cell surface glycoprotein CD4; Immunoglobulin homology  
 C/Keywords: Glycoprotein; membrane protein; surface antigen  
 F/219-300/Domain: immunoglobulin homology <IMM>

Query Match 19.1%; Score 508.5; DB 2; Length 457;  
 Best Local Similarity 49.3%; Pred. No. 9.2e-20;  
 Matches 104; Conservative 34; Mismatches 70; Indels 3; Gaps 2;

QY 1 MNRGVPFRHL--LLVQLALLPAPATQKRVVGGKGDVETLTCTASQKSIQPHMKNSQ 58  
 DB 1 MCRGSPFRHLPLLLQLSLVLTQKTVVAGKEGSELEPCESISRASAFMKSSDQ 60  
 QY 59 IKILGNGSFYTKPESKLNDRADSRRLWDQGNPLIILNKLKIEDSTTYICEVEDQKEEY 118  
 DB 61 KTLIGYKNKLILGSELYSRFDSRKNAWERSFPILIKLRNEDSQTYVCELENKEEY 120  
 QY 119 QLVFGLTANSPTHLQOGSLTITLES-PPGSSPSVQCRSPGKNIQGGKTLVSQLEIQ 177  
 DB 121 ELWVFVTFNPGFRLLQOGSLTILIDSNPKVSPDPIECGKSNIVKDSKAFSTHSLRIQ 180  
 QY 178 DSGTWTCTVQNGKVEFKIDIVVLAFOKAS 208  
 DB 181 DSGTWTCTVQNGKVEFKIDIVVLAFOKAS 211

## RESULT 10

S40343  
 Ig kappa chain V-J region - human  
 C/Species: Homo sapiens (man)  
 C/Date: 19-May-1994 #sequence revision 26-May-1995 #text\_change 21-Jan-2000  
 C/Accession: S40343  
 R/Klein, R.; Jaenichen, R.; Zachau, H.G.  
 Eur. J. Immunol. 23, 3248-3271, 1993  
 A>Title: Expressed human immunoglobulin chi genes and their hypermutation.  
 A/Reference number: S40343; MUID:94080891; PMID:8258341  
 A/Accession: S40343  
 A/Status: preliminary; translation not shown  
 A/Molecule type: mRNA  
 A/Residues: 1-128 <KLE>  
 A/Cross-references: EMBL:X72453; NID:g441374; PIDD:CAA51121.1; PIDD:g441375  
 C/Superfamily: Immunoglobulin V region; Immunoglobulin homology  
 C/Keywords: heterotetramer; immunoglobulin  
 F/36-110/Domain: immunoglobulin homology <IMM>

Query Match 19.0%; Score 507; DB 2; Length 128;  
 Best Local Similarity 89.0%; Pred. No. 3.3e-20;  
 Matches 97; Conservative 7; Mismatches 3; Indels 2; Gaps 1;

QY 387 ELELTQSPATLSVSPERATLSCRASESYSDDLAWYQKPGQAPRLIYGASTRATGVA 446  
 DB 21 EIVMTQSPATLSVSPERATLSCRAQSQSVSNLAWYQKPGQAPRLIYGASTRATGVA 80  
 QY 447 RFGSGSGAEFTLTSSLOSEPFAVYCCQINNMPRYTFGQTRLEIK 495

Db 81 RFGSGSGTEFTLTITSSLSQSEDFAVVYCCQYNNMP--YTFQGGTKLEIK 127

# RESULT 11

IG kappa chain V region - human

C:Species: Homo sapiens (man)  
C>Date: 02-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 21-Jan-2000  
C:Accession: S34005; S30524  
R:Marlette, X.; Tsapis, A.; Brucet, J.C.  
Eur. J. Immunol. 23, 846-851, 1993  
A:Title: Nucleotide sequence analysis of the variable domains of four human monoclonal A:Reference number: S34001; MUID:93209281; PMID:7681398  
A:Accession: S34005  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-107 <MAR>  
C:Cross-references: EMBL:218330  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
F:16-90/Domain: immunoglobulin homology <IMM>

Query Match 18.7%; Score 498; DB 2; Length 107;  
Best Local Similarity 88.1%; Pred. No. 8.2e-20;  
Matches 96; Conservative 7; Mismatches 4; Indels 2; Gaps 1;

Qy 387 ELETSPATLVSPPGERATISCRASBSVSDLMWYQKPGQARLLIYGASTATGVP 446

Db 1 EIVWTQSPATLVSPPGERATISCRASQSVSNLWYQKPGQARLLIYGASTATGIPA 60

Qy 447 RFGSGSGAETFTLTITSSLSQSEDFAVVYCCQYNNMPRYTGGTRLEIK 495  
Db 61 RFGSGSGTEFTLTITSSLSQSEDFAVVYCCQYNNMP--RTFGGTRLEIK 107

## RESULT 12

IG kappa chain - human

C:Species: Homo sapiens (man)  
C>Date: 06-Mar-1994 #sequence\_revision 26-May-1995 #text\_change 21-Jan-2000  
C:Accession: S40362  
R:Klein, R.; Jaenichen, R.; Zachau, H.G.  
Eur. J. Immunol. 23, 3248-3271, 1993  
A:Title: Expressed human immunoglobulin chi genes and their hypermutation.  
A:Reference number: S40312; MUID:94080891; PMID:8258341  
A:Accession: S40362  
A:Status: preliminary; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-117 <KLE>  
A:Cross-references: EMBL:X72472; NID:g441412; PID:g441413  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
F:26-100/Domain: immunoglobulin homology <IMM>

Query Match 18.7%; Score 498; DB 2; Length 117;  
Best Local Similarity 88.1%; Pred. No. 8.9e-20;  
Matches 96; Conservative 7; Mismatches 4; Indels 2; Gaps 1;

Qy 387 ELETSPATLVSPPGERATISCRASBSVSDLMWYQKPGQARLLIYGASTATGVP 446

Db 11 EIVWTQSPATLVSPPGERATISCRASQSVSNLWYQKPGQARLLIYGASTATGIPA 70

Qy 447 RFGSGSGAETFTLTITSSLSQSEDFAVVYCCQYNNMPRYTGGTRLEIK 495  
Db 71 RFGSGSGTEFTLTITSSLSQSEDFAVVYCCQYNNMP--LTFGGTRLEIK 117

## RESULT 13

RMWS14 T-cell surface glycoprotein CD4 precursor - mouse

N:Alternate names: T-cell differentiation antigen L3T4; T-cell surface antigen T4/Leu 3  
C:Species: Mus musculus (house mouse)

C>Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 09-Jul-2004  
C:Accession: A02110; A26038; A39893; A39955; I54564; I69018; A47642  
R:Tourville, B.; Gorman, S.D.; Field, E.H.; Hunkapiller, T.; Parnes, J.R.  
Science 234, 610-614, 1986  
A:Title: Isolation and sequence of L3T4 complementary DNA clones: expression in T cells  
A:Reference number: A02110; MUID:87018845; PMID:3094146  
A:Accession: A02110  
A:Molecule type: mRNA  
A:Residues: 1-457 <TOU>

A:Cross-references: UNIPROT:P06332; GB:M13816; NID:g192070; PIDN:AAA37267.1; PID:g309112  
R:Litman, D.R.; Gellner, S.N.  
Nature 325, 453-455, 1987  
A:Title: Unusual intron in the immunoglobulin domain of the newly isolated murine CD4 (L3T4) gene.

A:Reference number: A26038; MUID:87115821; PMID:302755  
A:Accession: A26038  
A:Molecule type: mRNA

A:Residues: 1-457 <LIT>  
A:Cross-references: GB:X04836; NID:g50353; PIDN:CAA28539.1; PID:g50354  
R:Gorman, S.D.; Tourville, B.; Parnes, J.R.  
Proc. Natl. Acad. Sci. U.S.A. 84, 7644-7648, 1987

A:Title: Structure of the mouse gene encoding CD4 and an unusual transcript in brain.  
A:Reference number: A39893; MUID:88041159; PMID:2823269  
A:Accession: A39893  
A:Molecule type: DNA

A:Residues: 1-25, 'E', '27-457 <GOR>  
A:Cross-references: GB:M17080; GB:J03003; NID:g192515; PIDN:AAA37402.1; PID:g387124  
R:Maddon, P.J.; Molineux, S.M.; Maddon, D.R.; Zimmermann, K.A.; Godfrey, M.; Alt, F.W.; (

Proc. Natl. Acad. Sci. U.S.A. 84, 9155-9159, 1987  
A:Title: Structure and expression of the human and mouse T4 genes.  
A:Reference number: A39955; MUID:88097446; PMID:3501122  
A:Accession: A39955  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA  
A:Residues: 25-457 <MAD>  
A:Note: the cited GenBank accession number, J03564, is not in release 101.0  
R:Parnes, J.R.; Hunkapiller, T.  
Immunol. Rev. 100, 109-127, 1987

A:Title: L3T4 and the immunoglobulin gene superfamily: New relationships between the immunoglobulin and the immunoglobulin gene superfamily.  
A:Reference number: I54564; MUID:88152875; PMID:3326818  
A:Accession: I54564  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA

A:Cross-references: GB:M36850; NID:g198670; PIDN:AAA39401.1; PID:g198671  
A:Accession: I69018  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA

A:Residues: 208-318 <RE2>  
A:Cross-references: GB:M36851; NID:g198672; PIDN:AAA39402.1; PID:g554183  
R:Classon, B.J.; Tsagaratos, J.; Kirschbaum, L.; Maddox, J.; Mackay, C.R.; Brandon, M.; McK  
Immunogenetics 23, 129-132, 1986

A:Title: The L3T4 antigen in mouse and the sheep equivalent are immunoglobulin-like.  
A:Reference number: A47642; MUID:8616694; PMID:3082751  
A:Accession: A47642  
A:Molecule type: protein  
A:Residues: 27-43 <CLA>

C:Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells of C:Genetics:  
A:Initons: 18/1; 74/1; 128/1; 207/1; 319/1; 386/1; 425/3; 448/2  
C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
C:Keywords: alternative initiator; duplication; glycoprotein; T-cell; transmembrane prot

F:1-26/Domain: signal sequence #status predicted <SIG>  
F:21-457/Product: CD4, brain-specific short form #status predicted <BRA>  
F:321-372/Domain: immunoglobulin homology <IM2>  
F:35-114/Domain: immunoglobulin homology <IM1>  
F:139-190/Domain: immunoglobulin homology #status atypical <IM2>  
F:220-301/Domain: immunoglobulin homology <IM3>

F:241-457/Product: CD4, brain-specific short form #status predicted <BRA>  
F:321-372/Domain: immunoglobulin homology <IM4>  
F:335-419/Domain: transmembrane #status predicted <TM>  
F:420-457/Domain: intracellular #status predicted <INT>  
F:42-112,159-188,328-370/Dsulfide bonds: #status predicted

F:187,298,323,392/Binding site: carbohydrate (Asn) (covalent) #status predicted



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: September 20, 2005, 12:03:07 ; Search time 137 Seconds  
(without alignments)  
1898.805 Million cell updates/sec

Title: US-09-936-702-3

Perfect score: 2668  
Sequence: 1 MNRGVPRHLLVLTQLALP.....GTRLEIKIVPRSGHHHHH 508

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1064	39.9	458	1	CD4_HUMAN
2	1040	39.0	458	1	CD4_PANTR
3	953	35.7	458	1	CD4_MACRA
4	951	35.6	458	1	CD4_MACRU
5	945	35.4	458	1	CD4_MACMU
6	944	35.4	458	1	CD4_MACNE
7	926	34.7	458	1	CD4_CERAB
8	824.5	30.9	244	2	065ZC8
9	820	30.7	397	1	CD4_CERPO
10	819	30.7	397	1	CD4_ERRPA
11	803	30.1	397	2	009261
12	803	30.1	397	2	009262
13	803	30.1	397	2	009263
14	803	30.1	397	2	009269
15	799	29.9	397	2	009259
16	799	29.9	397	2	009260
17	778.5	29.2	457	2	08H2R8
18	777.5	29.1	457	2	08H2R7
19	774.5	29.0	457	1	CD4_SAI5C
20	766	28.7	255	2	06KB05
21	760.5	28.5	240	2	065ZC9
22	751.5	28.2	487	2	065ZL2
23	750	28.1	487	2	07TQM2
24	734	27.5	293	2	09QYR0
25	685.5	25.7	241	2	0921A6
26	651	24.4	248	2	065Z07
27	632	23.7	463	1	CD4_CANFA
28	630	23.6	455	2	0710E2
29	625	23.4	455	2	09X578
30	618.5	23.2	459	1	CD4_RABIT
31	612	22.9	218	2	0925S1

32	596	22.3	432	2	06LBK1	Q6LBK1 canis fami
33	590.5	22.1	474	2	P79355	P79355 felis silve
34	585	21.9	457	2	06GYR3	06GYR3 sus scrofa
35	563	21.1	406	2	06R3N3	06R3N3 sus scrofa
36	563	21.1	447	2	06R3N4	06R3N4 sus scrofa
37	508.5	19.1	457	1	CD4_RAT	P05540 ratius norv
38	505.5	18.9	129	1	KV3H_HUMAN	P04207 homo sapien
39	498	18.7	457	1	CD4_MOUSE	P06332 mus musculu
40	490	18.4	457	2	061396	061396 mus musculu
41	485.5	18.2	109	2	09UL85	09UL85 homo sapien
42	480.5	18.0	109	1	KV3P_HUMAN	P01624 homo sapien
43	480	18.0	108	2	09UL83	09UL83 homo sapien
44	472.5	17.7	205	2	06GMW0	06GMW0 homo sapien
45	459.5	17.2	433	2	055054	055054 mus musculu

## ALIGNMENTS

RESULT 1	CD4_HUMAN	STANDARD;	PRT;	458 AA.
AC	P01730;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-NOV-1988 (Rel. 09, Last sequence update)			
DT	25-JAN-2005 (Rel. 46, Last annotation update)			
DE	T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).			
GN	Name=CD4;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85254948; PubMed=2990730;			
RA	Maddon P.J., Littman D.R., Godfrey M., Maddon D.E., Chess L., Axel R.;			
RT	"The isolation and nucleotide sequence of a cDNA encoding the T cell surface protein T4: a new member of the immunoglobulin gene family.";			
RL	Cell 42:93-104(1985).			
RN	[2]			
RP	REVISION TO 26.			
RX	MEDLINE=89028665; PubMed=3263213; DOI=10.1016/0092-8674(88)90211-5;			
RA	Littman D.R., Maddon P.J., Axel R.;			
RT	"Corrected CD4 sequence.";			
RL	Cell 55:541-541(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96303695; PubMed=8723724;			
RA	Ansari-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E., Spanos S.,			
RT	"A gene-rich cluster between the CD4 and triosephosphate isomerase genes at human chromosome 12p13.";			
RL	Genome Res. 6:314-326(1996).			
RN	[4]			
RP	SEQUENCE FROM N.A., AND VARIANT TRP-265.			
RX	MEDLINE=91216786; PubMed=1708753; DOI=10.1016/0198-8859(91)90077-M;			
RA	Hodge T.W., Saaso D.R., McDougal J.S.;			
RT	"Humans with OKT4-epitope deficiency have a single nucleotide base change in the CD4 gene, resulting in substitution of TRP240 for ARG240.";			
RL	Hum. Immunol. 30:99-104(1991).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Pancreas;			
RX	MEDLINE=92388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Krausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,			
RA	Stausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Altschul S.F., Zeeberg B., Buertow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatcenco L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Scapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,			

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., McWiley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fley J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalka U., Smillie D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RL [6]  
 RN SEQUENCE OF 28-424 FROM N.A.  
 RP TISSUE=Blood;  
 RC MEDLINE=9049640; PubMed=1425921;  
 RX Fomsgaard A., Hirsch V.M., Johnson P.R.,  
 RA "Cloning and sequences of primate CD4 molecules: diversity of the  
 RT cellular receptor for simian immunodeficiency virus/human  
 RT Immunodeficiency virus.";  
 RL Eur. J. Immunol. 22:2973-2981(1992).  
 RN [7]  
 RP SEQUENCE OF 26-394 AND DISULFIDE BOND.  
 RX MEDLINE=9078232; PubMed=2592374;  
 RA Carr S.A., Hemling M.E., Folena-Waserman G., Sweet R.W., Anumula K.,  
 RA Barr J.R., Huddleston M.J., Taylor P.,  
 RT "Protein and carbohydrate structural analysis of a recombinant soluble  
 RT CD4 receptor by mass spectrometry.";  
 RL J. Biol. Chem. 264:21286-21295(1989).  
 RN [8]  
 RP SEQUENCE OF 26-40.  
 RX PubMed=15340161; DOI=10.1110/ps.04682504;  
 RA Zhang Z., Henzel W.J.,  
 RT "Signal peptide prediction based on analysis of experimentally  
 RT verified cleavage sites.";  
 RL Protein Sci. 13:2819-2824(2004).  
 RN [9]  
 RP REMOVAL FROM CELL SURFACE BY HIV-1 NEF, AND MUTAGENESIS OF MET-432;  
 RP SER-433; 438-LEU-LEU-439 AND SER-440.  
 RX PubMed=8124721;  
 RA Alken C., Konner J., Landau N.R., Lenburg M.E., Trono D.,  
 RT "Nef induces CD4 endocytosis: requirement for a critical dileucine  
 RT motif in the membrane-proximal CD4 cytoplasmic domain.";  
 RL Cell 76:853-864(1994).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 26-208.  
 RX MEDLINE=81061881; PubMed=1701030; DOI=10.1038/348411a0;  
 RA Wang J., Yan Y., Garret T.P., Liu J., Rodgers D.W., Garlick R.L.,  
 RA Tarr G.E., Husain Y., Reinherz E.L., Harrison S.C.,  
 RT "Atomic structure of a fragment of human CD4 containing two  
 RT immunoglobulin-like domains.";  
 RL Nature 348:411-418(1990).  
 RN [11]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 26-208.  
 RX MEDLINE=1061882; PubMed=2247146; DOI=10.1038/348419a0;  
 RA Ryu S.-E., Kwong P.D., Truneh A., Porter T.G., Atchos J.,  
 RA Rosenberg M., Dai X., Xuong N.-H., Axel R., Sweet R.W.,  
 RA Hendrickson W.A.,  
 RT "Crystal structure of an HIV-binding recombinant fragment of human  
 RT CD4.";  
 RL Nature 348:419-426(1990).  
 RN [12]  
 RP X-RAY CRYSTALLOGRAPHY (3.9 ANGSTROMS) OF 26-388.  
 RX MEDLINE=97311402; PubMed=9168119;  
 RA Wu H., Kwong P.D., Hendrickson W.A.,  
 RT "Dimeric association and segmental variability in the structure of  
 RT human CD4.";  
 RL Nature 387:527-530(1997).  
 RN [13]  
 RP PALMITOYLATION.  
 RX MEDLINE=92317088; PubMed=1618861;  
 RA Crise B., Rose J.K.,

RT "Identification of palmitoylation sites on CD4, the human  
 RT Immunodeficiency virus receptor.";  
 RL J. Biol. Chem. 267:13593-13597(1992).  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Removed from plasma  
 CC membrane by HIV-1 Nef protein that increases clathrin-dependent  
 CC endocytosis of this antigen to target it to lysosomal degradation.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -1- DATABASE: NAME=PROIM; NOTE=CD guide CD4 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/blast/blast.cgi?cd4".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by, and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M12807; AAA35572.1; -  
 DR EMBL: U47924; AAB51309.1; -  
 DR EMBL: M35160; AAL16069.1; -  
 DR EMBL: BC025782; AAL25782.1; -  
 DR PIR: A90872; RWHUT4.  
 DR PDB: 1CDH; X-ray; @=26-203.  
 DR PDB: 1CDI; X-ray; @=25-203.  
 DR PDB: 1CDJ; X-ray; @=26-203.  
 DR PDB: 1CDU; X-ray; @=26-203.  
 DR PDB: 1CDV; X-ray; @=26-203.  
 DR PDB: 1G9M; X-ray; C=26-210.  
 DR PDB: 1G9N; X-ray; C=26-210.  
 DR PDB: 1GCI; X-ray; C=26-210.  
 DR PDB: 1JL4; X-ray; D=26-203.  
 DR PDB: 1Q68; NMR; A=421-458.  
 DR PDB: 1WER; NMR; @=427-445.  
 DR PDB: 1WIO; X-ray; A/B=26-388.  
 DR PDB: 1WIP; X-ray; A/B=26-388.  
 DR PDB: 1WIO; X-ray; A/B=26-388.  
 DR PDB: 3CD4; X-ray; @=26-207.  
 DR GlycoSuiteDB: P01730; -  
 DR GeneW: HGNC:1678; CD4.  
 DR H-InvDB: HIX0023001; -  
 DR MIM: 186940; -  
 DR GO: GO:0005886; C:Plasma membrane; TAS.  
 DR GO: GO:0042101; C:T-cell receptor complex; NAS.  
 DR GO: GO:0015026; F:coreceptor activity; NAS.  
 DR GO: GO:0015029; F:internalization receptor activity; TAS.  
 DR GO: GO:0042289; F:MHC class II protein binding; NAS.  
 DR GO: GO:0004888; F:transmembrane receptor activity; TAS.  
 DR GO: GO:0006955; P:immune response; NAS.  
 DR GO: GO:0045085; P:positive regulation of interleukin-2 biosyn. .; NAS.  
 DR GO: GO:0030217; P:T-cell differentiation; NAS.  
 DR GO: GO:0045058; P:T-cell selection; NAS.  
 DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. .; NAS.  
 DR InterPro: IPR008424; CD2.  
 DR InterPro: IPR000973; CD4 TCAG.  
 DR InterPro: IPR007110; Ig-Tlike.  
 DR InterPro: IPR003596; Ig-V.  
 DR Pfam: PF00047; Ig\_2.  
 DR PRINTS: PR00692; CD4TCANTIGEN.  
 DR SMART: SM00406; IGV\_1.  
 DR PROSITE: PS50835; IG\_LIKE\_1.  
 KW 3D-structure; Direct protein sequencing; Glycoprotein;  
 KW Immune response; Immunoglobulin domain; Lipoprotein; Palmitate;  
 KW Polymorphism; Repeat; Signal; T-cell; Transmembrane.  
 FT STGNAL 1 25  
 FT CHAIN 26 458  
 FT DOMAIN 26 396  
 FT TRANSMEM 397 418  
 FT DOMAIN 419 458  
 FT T-cell surface glycoprotein CD4.  
 FT Extracellular (Potential).  
 FT Potential.  
 FT Cytoplasmic (Potential).

```

FT DOMAIN 26 125 Ig-like V-type.
FT DOMAIN 126 203 Ig-like C2-type 1.
Query Match 39.9%; Score 1064; DB 1; Length 458;
Best Local Similarity 100.0%; Pred. No. 1.9e-59;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPPATQGGKVVGLKGGDTVELTCFASOKKSIOFHMKNSQIK 60
DB 1 MNRGVPFRHLVLVQLALPPATQGGKVVGLKGGDTVELTCFASOKKSIOFHMKNSQIK 60
QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGGSLTLTLESPGSSPSVQCRSPRKNIOGKTLVSQLELDG 180
DB 121 LVFGLTANSDTHLLQGGSLTLTLESPGSSPSVQCRSPRKNIOGKTLVSQLELDG 180
QY 181 TWCTVLOKQKVEFKIDIVLAFOKAS 208
DB 181 TWCTVLOKQKVEFKIDIVLAFOKAS 208

RESULT 2
CD4_PANTR STANDARD; PRT: 458 AA.
AC P16004;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).
GN Name=CD4;
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
ON NCBI_TaxId=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90182664; PubMed=2107024; DOI=10.1016/0092-8674(90)90089-W;
RA Cameron D., Seed B.;
RT "A CD4 domain important for HIV-mediated syncytium formation lies
RT outside the virus binding site."
RT Cell 60:747-754(1990).
RN [2]
RP SEQUENCE OF 26-424 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus."
RT Eur. J. Immunol. 22:2973-2981(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL: M3135; AAA35407.1; -
CC DR EMBL: X73323; CAAS1749.1; -
CC DR PIR: B32722; RWC274.
CC DR HSSP: P01730; IWIO.

```

```

DR GO: GO:0042101; C:T-cell receptor complex; ISS.
DR GO: GO:0015026; F:coreceptor activity; ISS.
DR GO: GO:0042289; F:MHC class II protein binding; ISS.
DR GO: GO:0006555; P:immune response; ISS.
DR GO: GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO: GO:0030217; P:T-cell differentiation; ISS.
DR GO: GO:0045058; P:T-cell selection; ISS.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR009973; CD4_TCSA.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig_2.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SM00406; IGv_1.
DR PROSITE: PSS0835; IG-LIKE_1.
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSMEM 397 418
FT DOMAIN 419 458
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 296 296
FT CARBOHYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT LIPID 419 419
FT LIPID 422 422
FT CONFLICT 42 42
FT CONFLICT 62 62
FT CONFLICT 191 191
SQ SEQUENCE 458 AA; 51057 MW; A7C3AC8A5257D3AD CRC64;

Query Match 39.0%; Score 1040; DB 1; Length 458;
Best Local Similarity 98.1%; Pred. No. 5.9e-58;
Matches 204; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPPATQGGKVVGLKGGDTVELTCFASOKKSIOFHMKNSQIK 60
DB 1 MNRGVPFRHLVLVQLALPPATQGGKVVGLKGGDTVELTCFASOKKSIOFHMKNSQIK 60
QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGGSLTLTLESPGSSPSVQCRSPRKNIOGKTLVSQLELDG 180
DB 121 LVFGLTANSDTHLLQGGSLTLTLESPGSSPSVQCRSPRKNIOGKTLVSQLELDG 180
QY 181 TWCTVLOKQKVEFKIDIVLAFOKAS 208
DB 181 TWCTVLOKQKVEFKIDIVLAFOKAS 208

RESULT 3
CD4_MACFA STANDARD; PRT: 458 AA.
AC P79185;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).
GN Name=CD4;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;

```

CC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RA Tatemura M., Yabe M., Yamada Y.K.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D63349; BAA09673.1; -.  
 DR HSSP; P01730; IMBR.  
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.  
 DR GO; GO:0006955; P:immune response; ISS.  
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.  
 DR GO; GO:0030217; P:T-cell differentiation; ISS.  
 DR GO; GO:0045058; P:T-cell selection; ISS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TCAG.  
 DR InterPro; IPR007110; Ig\_Like.  
 DR InterPro; IPR003596; Ig\_v.  
 DR Pfam; PF00047; Ig\_2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KM Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;  
 KM Palmitate; Repeat; Signal; T-cell; Transmembrane.  
 FT STGNAL 1 25  
 FT CHAIN 26 458  
 FT DOMAIN 26 396  
 FT TRANSMEM 397 418  
 FT DOMAIN 419 458  
 FT DOMAIN 126 125  
 FT DOMAIN 203 317  
 FT DOMAIN 318 374  
 FT DOMAIN 42 42  
 FT CARBOHYD 296 296  
 FT CARBOHYD 325 325  
 FT DISULFID 41 109  
 FT DISULFID 155 184  
 FT DISULFID 328 370  
 FT LIPID 419 419  
 FT LIPID 422 422  
 SQ SEQUENCE 458 AA; 50872 MW; 9105479FB5C56FF7 CRC64;

Query Match 35.7%; Score 953; DB 1; Length 458;  
 Best Local Similarity 88.9%; Pred. No. 1.9e-52;  
 Matches 185; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 121 LVFGLTANSDTHLLOQSITLTLESPGSSPSVQCRSPGRKNIQGGKTLVSQLEIQDSG 180  
 DB 121 LVFGLTANSDTHLLEQSITLTLESPGSSPSVKRSPGKNIQGGRTLSVPQLERQDSG 180  
 QY 181 TWTCYTLQNKQKVEPKIDIVLAFQKAS 208  
 DB 181 TWTCVSDQKQVEFKIDIVLAFQKAS 208  
 RESULT 4  
 CD4\_MACEFU STANDARD; PRT; 458 AA.  
 AC P79184;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen  
 DE T/Len-3).  
 GN Name=CD4;  
 OS Macaca fuscata fuscata (Japanese macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9543;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hashimoto O., Tatemura M.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D63348; BAA09672.1; -.  
 DR HSSP; P01730; IMBR.  
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.  
 DR GO; GO:0006955; P:immune response; ISS.  
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.  
 DR GO; GO:0030217; P:T-cell differentiation; ISS.  
 DR GO; GO:0045058; P:T-cell selection; ISS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TCAG.  
 DR InterPro; IPR007110; Ig\_Like.  
 DR InterPro; IPR003596; Ig\_v.  
 DR Pfam; PF00047; Ig\_2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KM Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;  
 KM Palmitate; Repeat; Signal; T-cell; Transmembrane.  
 FT STGNAL 1 25  
 FT CHAIN 26 458  
 FT DOMAIN 26 396  
 FT TRANSMEM 397 418  
 FT DOMAIN 419 458  
 FT DOMAIN 126 125  
 FT DOMAIN 203 317  
 FT DOMAIN 318 374  
 FT CARBOHYD 42 42



```

FT CARBOHYD 296 296 N-linked (GlcNAc... ) (By similarity).
FT CARBOHYD 325 325 N-linked (GlcNAc... ) (By similarity).
FT DISULFID 41 109 By similarity.
FT DISULFID 155 184 By similarity.
FT DISULFID 328 370 By similarity.
FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
SQ SEQUENCE 458 AA; 50828 MW; 76B3E7EF08185535 CRC64;

Query Match 35.4%; Score 951; DB 1; Length 458;
Best Local Similarity 88.0%; Pred. No. 2,6e-52;
Matches 184; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLIVLQALALPAATGKRVLGKGGDYELTCTASOKKSIOFHWKNSNOIK 60
D 1 MNRGIPFRHLIVLQALLPAVVGKRVVGLCKKDDYELTCNASKKNTQFHWKNSNOIK 60
QY 61 ILNGSGFLTKGSPKSLNDRADSRRLMDQGNFLLINKLKIEDSDTYICEVEDQKEVOL 120
D 61 ILGIQSGFLTKGSPKSLNDRADSRRLMDQGNFLLINKLKIEDSDTYICEVENKEVEL 120
QY 121 LVFGLTANSTHLLQGSGLTILTESPPGSSPVCCSRPRGNKIQGKTLISVQLELDDSG 180
D 121 LVFGLTANSTHLLQGSGLTILTESPPGSSPVCCSRPRGNKIQGKTLISVQLELDDSG 180
QY 181 TWICTVYQNKQKVEFKIDIVVLAFOKAS 208
D 181 TWICTVYQNKQKVEFKIDIVVLAFOKAS 208
D 181 TWICTVYQNKQKVEFKIDIVVLAFOKAS 208

RESULT 5
CD4 - MACMU STANDARD; PRT; 458 AA.
ID CD4 - MACMU STANDARD; PRT; 458 AA.
AC P16T03; Q29617;
DT 01-APR-1990 (Rel. 14, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN Name=CD4;
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90182664; PubMed=2107024; DOI=10.1016/0092-8674(90)90089-W;
RA Camerini D., Seed B.;
RT "A CD4 domain important for HIV-mediated syncytium formation lies
RT outside the virus binding site.";
RL Cell 60:747-754(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymocytes;
RA Hashimoto O., Tatsumi M.;
RT "Molecular cloning and expression of macaque CD4s.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomesgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus.";
RL Eur. J. Immunol. 22:2973-2981(1992).
RN [4]
RP SEQUENCE OF 107-192 FROM N.A.
RX MEDLINE=98320644; PubMed=3656488;
RA Harris E.B., Disotell T.R.;
RT "Nuclear gene trees and the phylogenetic relationships of the
RT mangabeys (Primates: Papionini).";

```

```

RL Mol. Biol. Evol. 15:892-900(1998).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p55-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M31134; AAA36838.1; -.
DR EMBL; D63347; BAA09671.1; -.
DR EMBL; X73326; CAA51752.1; -.
DR EMBL; AF057385; AAC25129.1; -.
DR HSP: P01730; 1MBR.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro; IPR008424; CD2.
DR InterPro; IPR000973; CD4_TcAg.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 25
FT CHAIN 26 458 T-cell surface glycoprotein CD4.
FT DOMAIN 26 396 Extracellular (Potential).
FT TRANSMEM 397 418 Potential.
FT DOMAIN 419 458 Cytoplasmic (Potential).
FT DOMAIN 26 125 Ig-like V-type.
FT DOMAIN 126 203 Ig-like C2-type 1.
FT DOMAIN 204 317 Ig-like C2-type 2.
FT DOMAIN 318 374 Ig-like C2-type 3.
FT CARBOHYD 296 296 N-linked (GlcNAc... ) (By similarity).
FT CARBOHYD 325 325 N-linked (GlcNAc... ) (By similarity).
FT DISULFID 41 109 By similarity.
FT DISULFID 155 184 By similarity.
FT DISULFID 328 370 By similarity.
FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
FT CONFLICT 42 42 N -> T (in Ref. 1).
FT CONFLICT 62 62 L -> S (in Ref. 3).
FT CONFLICT 67 67 L -> S (in Ref. 2).
FT CONFLICT 169 169 I -> L (in Ref. 2).
FT CONFLICT 191 191 K -> N (in Ref. 2).
FT CONFLICT 248 248 S -> P (in Ref. 2).
FT CONFLICT 265 265 R -> Q (in Ref. 3).
FT CONFLICT 349 349 A -> T (in Ref. 2).
SQ SEQUENCE 458 AA; 50884 MW; 8BB80339FAFEC08 CRC64;

Query Match 35.4%; Score 945; DB 1; Length 458;
Best Local Similarity 88.0%; Pred. No. 6,2e-52;
Matches 183; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLIVLQALALPAATGKRVLGKGGDYELTCTASOKKSIOFHWKNSNOIK 60
D 1 MNRGIPFRHLIVLQALLPAVVGKRVVGLCKKDDYELTCNASKKNTQFHWKNSNOIK 60

```

```

QY 61 ILNGQSFLLTKGSKLNDRAISRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGIQGFLLTKGSKLSDRAISRSLMDQGNFPLIINKLKIEDSDTYICEVENKEEVEL 120
QY 121 LVFGLTANSTHLLQOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLEODSG 180
DB 121 LVFGLTANSTHLLQOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLEODSG 180
QY 181 TWTCVQLQNKQKVEFKIDIVLAFOKAS 208
DB 181 TWTCVQLQNKQKVEFKIDIVLAFOKAS 208

RESULT 6
CD4_MACNE STANDARD; PRT; 458 AA.
ID 008340: P79196;
AC 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN Name=CD4;
OS Macaca nemestrina (Pig-tailed macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
NCBI_TaxID=9545;
[1]
RN SEQUENCE FROM N.A.
RA Hashimoto O., Tatsumi M.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP TISSUE=Blood;
RC MEDLINE=93049640; PubMed=1425921;
RA Fomsgard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus.";
RL Eur. J. Immunol. 23:2973-2981(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (by similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D63346; BAA09670.1; -
CC EMBL; X73325; CAA51751.1; -
CC HSSP; P01730; 1WBR.
DR GO: GO:0042101; C:T-cell receptor complex; ISS.
DR GO: GO:0015026; F:coreceptor activity; ISS.
DR GO: GO:0042289; F:MHC class II protein binding; ISS.
DR GO: GO:0006955; P:immune response; ISS.
DR GO: GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO: GO:0030217; P:T-cell differentiation; ISS.
DR GO: GO:0045058; P:T-cell selection; ISS.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR000973; CD4_TcRA.
DR InterPro: IPR007110; Ig_Like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.

```

```

DR SMART; SMO0406; IGV; 1.
DR PROSITE; PS50835; IG_Like; 1.
KM Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KM Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGML 1 25
FT CHAIN 26 458
FT TRANSMEM 26 396
FT TRANSMEM 397 418
FT DOMAIN 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 42 42
FT CARBOHYD 296 296
FT CARBOHYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT DISULFID 419 419
FT LIPID 422 422
FT LIPID 57 57
FT CONFLICT 91 105
FT CONFLICT 105 105
FT CONFLICT 113 113
FT CONFLICT 302 302
FT CONFLICT 349 349
SQ SEQUENCE 458 AA; 50905 MW; 751A9BA2C8B3EB16 CRC64;

Query Match 35.4%; Score 944; DB 1; Length 458;
Best Local Similarity 88.0%; Pred. No. 7, 2e-52;
Matches 183; Conservative 14; Mismatches 11; Indels 0; Gaps 0;

QY 1 MNRGVEFRLHLVQLALPAATQGRKVVLGKKGVETVETCTASQKKSIOFHKNSNQIK 60
DB 1 MNRGVEFRLHLVQLALPAATQGRKVVLGKKGVETVETCTASQKKSIOFHKNSNQIK 60
QY 61 ILNGQSFLLTKGSKLNDRAISRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGIQGFLLTKGSKLSDRAISRSLMDQGNFPLIINKLKIEDSDTYICEVENKEEVEL 120
QY 121 LVFGLTANSTHLLQOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLEODSG 180
DB 121 LVFGLTANSTHLLQOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLEODSG 180
QY 181 TWTCVQLQNKQKVEFKIDIVLAFOKAS 208
DB 181 TWTCVQLQNKQKVEFKIDIVLAFOKAS 208

RESULT 7
CD4_CERAE STANDARD; PRT; 458 AA.
ID 008338: O02805; 077593; 028217;
AC 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 29-MAR-2004 (Rel. 43, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN Name=CD4;
OS Cercopithecus aethiops (Green monkey) (Griwet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
NCBI_TaxID=9534;
[1]
RN SEQUENCE FROM N.A.
RA Hashimoto O., Tatsumi M.;
RL "Molecular cloning and expression of african green monkey CD4.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP TISSUE=Blood;
RC

```

RX MEDLINE=93049640; Pubmed=1425921;  
 RA Fomsgaard A., Hirsch V.M., Johnson P.R.;  
 RT "Cloning and sequences of primate CD4 molecules: diversity of the  
 RT cellular receptor for simian immunodeficiency virus/human  
 RT immunodeficiency virus";  
 RL Eur. J. Immunol. 22:2973-2981(1992).  
 RN [3]  
 RP SEQUENCE OF 28-424 FROM N.A.  
 RC TISSUE=Peripheral blood;  
 RX MEDLINE=98017879; Pubmed=9379478;  
 RA Fomsgaard A., Mueller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,  
 RA Corbet S., Barre-Sinoussi F., Allan J.S.;  
 RT Relation between phylogeny of African green monkey CD4 genes and  
 RT their respective simian immunodeficiency virus genes";  
 RL J. Med. Primatol. 26:120-128(1997).  
 RN [4]  
 RP SEQUENCE OF 107-192 FROM N.A.  
 RX MEDLINE=98320644; Pubmed=9656488;  
 RA Harris E.E., Disotell T.R.;  
 RT "Nuclear gene trees and the phylogenetic relationships of the  
 RT mangabeys (Primates: Papionini).";  
 RL Mol. Biol. Evol. 15:892-900(1998).  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D86589; BAAL132.1; -;  
 DR EMBL; X73323; CAAS1748.1; -;  
 DR EMBL; AF001226; AAB60873.1; -;  
 DR EMBL; AF001228; AAB60875.1; -;  
 DR EMBL; AF057380; AAC25124.1; -;  
 DR HSSP; P01703; IWIQ.  
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TcAg.  
 DR InterPro; IPR007110; IG-1Ike.  
 DR Pfam; PF00047; Ig; 2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;  
 KW Palmitate; Repeat; Signal; T-cell; Transmembrane.  
 FT SIGNL 1 25  
 FT CHAIN 26 458  
 FT DOMAIN 26 396  
 FT TRANSMEM 397 418  
 FT DOMAIN 419 458  
 FT DOMAIN 26 125  
 FT DOMAIN 126 203  
 FT DOMAIN 204 317  
 FT DOMAIN 318 374  
 FT CARBOHYD 42 42  
 FT CARBOHYD 281 281  
 FT CARBOHYD 296 296  
 FT CARBOHYD 325 325  
 FT DISULFID 41 109

FT DISULFID 155 184 By similarity.  
 FT DISULFID 328 370 By similarity.  
 FT LIPID 419 419 S-palmitoyl cysteine (By similarity).  
 FT LIPID 422 422 S-palmitoyl cysteine (By similarity).  
 FT CONFLICT 46 46 K -> N (in Ref. 2 and 3; AAB60875).  
 FT CONFLICT 46 46 I -> T (in Ref. 3; AAB60873).  
 FT CONFLICT 59 59 I -> T (in Ref. 1).  
 FT CONFLICT 115 115 K -> E (in Ref. 1).  
 FT CONFLICT 165 165 G -> V (in Ref. 3; AAB60873 and 4).  
 FT CONFLICT 200 200 M -> V (in Ref. 2 and 3).  
 FT CONFLICT 227 227 F -> L (in Ref. 3; AAB60873).  
 FT CONFLICT 271 271 K -> E (in Ref. 3; AAB60873).  
 FT CONFLICT 281 281 N -> H (in Ref. 3; AAB60873).  
 SQ SEQUENCE 458 AA; 51158 MW; FCS23D2EDD1F72E7 CRC64;  
 Query Match 34.7%; Score 926; DB 1; Length 458;  
 Best Local Similarity 86.5%; Pred. No. 1e-50;  
 Matches 180; Conservative 12; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 MNRGVEFRHLVLQLALPAATQGGKVVLGKGGDTVELCTASQKKSIOFHKNSNOIK 60  
 DB 1 MMNGIFRHLVLQLALPAATQGGKVVLGKGGDTVELCTASQKKTTFQHKNSNOIK 60  
 QY 61 ILGNQGSFLTKGSPKLNDRPADSRRLMDQGNPLIITKIKEDSDTYICEVEDQKEVOL 120  
 DB 61 ILGKQGSFLTKGSSKLRDRIDSRKSLMDQGSFMIITKIKEDSDTYICEVENKKEEVL 120  
 QY 121 LVFGLRANSDFTHLQGSILTLTLESPGSSPSVQCSPPRGKNIQGGKTLSVQLQLDSG 180  
 DB 121 LVFGLRANSDFTHLQGSILTLTLESPGSSPSVQCSPPRGKNIQGGKTLSVQLQLDSG 180  
 QY 181 TWTCVTLQNGKVEFKIDIVLAFQKAS 208  
 DB 181 TWTCVTLQNGKVEFKIDIVLAFQKAS 208  
 QY 181 TWTCVTLQNGKVEFKIDIVLAFQKAS 208  
 DB 181 TWTCVTLQNGKVEFKIDIVLAFQKAS 208  
 RESULT 8  
 Q65ZC8 PRELIMINARY; PRT; 244 AA.  
 ID 065ZC8;  
 AC 25-OCT-2004 (T-EMBLrel. 28, Created)  
 DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)  
 DE Single-chain Fv (Fragment).  
 GN Name=scFv;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97362799; Pubmed=9219263;  
 RA Kontermann R.E., Wing M.G., Winter G.;  
 RT "Complement recruitment using bispecific diabodies.";  
 RL Nat. Biotechnol. 15:629-631(1997).  
 DR EMBL; Y13057; CAA73500.1; -;  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR007110; Ig-1Ike.  
 DR InterPro; IPR003596; Ig-V.  
 DR Pfam; PF00047; Ig; 2.  
 DR SMART; SM00409; IGV; 2.  
 DR SMART; SM00406; IGV; 2.  
 DR PROSITE; PS50835; IG\_LIKE; 2.  
 FT NON\_TER 1 244  
 FT NON\_TER 244 244  
 SQ SEQUENCE 244 AA; 26127 MW; 4B1F17868338F2BF CRC64;  
 Query Match 30.9%; Score 824.5; DB 2; Length 244;  
 Best Local Similarity 64.7%; Pred. No. 1.3e-44;  
 Matches 163; Conservative 29; Mismatches 51; Indels 9; Gaps 2;  
 QY 244 QVQLLESGAEVKKPGSSSVVSCASGDTFIRYSFTWVROAPQGLBWKGRITITILDVAHY 303  
 DB 1 QVQLVDSGAEVKKPGSSSVVSCASGDTFIRYSFTWVROAPQGLBWKGRITITILDVAHY 303

QY 304 APHIGRVTITADKSTSTVYLRLNLRLSDTAVYFCAGVYEGEADGEYDNNGLKHMGO 363  
 DB 61 AGPFQGVMTWTRDTSTLSAAVMEVSRRLRSDTAVYCARREGTGAIVG-----MDVWGQ 113  
 QY 364 GTLVVYTSGGGGGGGGGGGGSELELTQSPATLSVSPGRATLSCASASVSDDLAWYQ 423  
 DB 114 GTLVVYTSGGGGGGGGGGGGSDIQMTQSPSTLSASIGRVITTCRASBGIVHMLAWYQ 173  
 QY 424 QKRGQAPRLIYAGSTRATGVPARFSGSGSGAEPTLTISLQSEDFAVYVYCOQYNNMPPR 483  
 DB 174 QKRGKAPKFIITYKASSLSASGAPSRFSGSGGTPTLTITSSLQPDPAFTYICQOYSNP-- 231  
 QY 484 YTFGGTRELRIK 495  
 DB 232 LTFGGGTKEIK 243

RESULT 9  
 CD4\_CERTO STANDARD; PRT; 397 AA.  
 ID\_008336;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Leu-3) (Fragment).  
 GN Name=CD4;  
 OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Cercopithecus.  
 NC NCBITaxID=9531;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RX MEDLINE=93049640; PubMed=1425921;  
 RA Fomsgaard A., Hirsch V.M., Johnson P.R.;  
 RT "Cloning and sequences of primate CD4 molecules: diversity of the  
 RT cellular receptor for simian immunodeficiency virus/human  
 RT immunodeficiency virus.";  
 RL Eur. J. Immunol. 22:2973-2981(1992).  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or  
 CC send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL: X73328; CAAS1754.1; -;  
 CC EMBL: X73327; CAAS1753.1; -;  
 CC HSPF: P01730; MWQ.  
 DR GO: GO:0042101; C:T-cell receptor complex; ISS.  
 DR GO: GO:0015026; F:coreceptor activity; ISS.  
 DR GO: GO:0042289; F:MHC class II protein binding; ISS.  
 DR GO: GO:0006955; P:immune response; ISS.  
 DR GO: GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.  
 DR GO: GO:0030217; P:T-cell differentiation; ISS.  
 DR GO: GO:0045058; P:T-cell selection; ISS.  
 DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.  
 DR InterPro: IPR008424; CD2.  
 DR InterPro: IPR000973; CD4 TCAG.  
 DR InterPro: IPR007110; Ig-Like.  
 DR InterPro: IPR003596; Ig\_v.  
 DR Pfam: PF00047; Ig; 2.

DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KM Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;  
 KM Palmitate; Repeat; T-cell; Transmembrane.  
 FT NON\_TER 1  
 FT DOMAIN 1  
 FT TRANSMEM 370 391 Extracellular (Potential).  
 FT POTENTIAL 392 397 Potential.  
 FT CYTOPLASMIC (Potential).  
 FT DOMAIN 392 397 Cytoplasmic (Potential).  
 FT DOMAIN 392 397 Ig-like V-type.  
 FT DOMAIN 392 397 Ig-like C2-type 1.  
 FT DOMAIN 392 397 Ig-like C2-type 2.  
 FT DOMAIN 392 397 Ig-like C2-type 3.  
 FT DOMAIN 392 397 N-linked (GlcNAc. . .) (Potential).  
 FT CARBOHYD 269 269 N-linked (GlcNAc. . .) (Potential).  
 FT CARBOHYD 298 298 N-linked (GlcNAc. . .) (Potential).  
 FT DISULFID 14 82 N-linked (GlcNAc. . .) (Potential).  
 FT DISULFID 128 157 By similarity.  
 FT DISULFID 301 343 By similarity.  
 FT LIPID 392 392 S-palmitoyl cysteine (By similarity).  
 FT LIPID 395 395 S-palmitoyl cysteine (By similarity).  
 FT VARIANT 20 20 Missing.  
 FT VARIANT 43 43 T -> I.  
 FT VARIANT 86 86 N -> D.  
 FT VARIANT 96 96 F -> L.  
 FT VARIANT 173 173 V -> M.  
 FT VARIANT 316 316 R -> K.  
 FT NON\_TER 397 397  
 SQ SEQUENCE 397 AA; 43926 MW; 8660B636D2DB38A7 CRC64;  
 Query Match 30.7%; Score 820; DB 1; Length 397;  
 Best Local Similarity 87.8%; Pred. No. 4,5e-44;  
 Matches 159; Conservative 11; Mismatches 11; Indels 0; Gaps 0;  
 QY 28 VVLGKKGVVELCTFSQKKSIOFHKNSNOIKIIGNSGFLTKGPKLNDRAADSRRLM 87  
 DB 1 VVLGKKGVVELCANQSKSIOFHKNSKOKIKILNQSGFLTKGSKLSDRAADSRRLM 60  
 QY 88 DQGNFPLIKNLKIEDSDTYICEVEDQKEVOLVGLTRANSDTHLQGSQSLTLTLESP 147  
 DB 61 DQCFSMIILKNLKEIDSETYICEVENKKEVELVGLTRANSDTHLEQSLTLTLESP 120  
 QY 148 GSSPSVQCSPPKKNIOGKTLVSQLELDGSGTWCTYLVQNKKEFKIDIVLAFQRA 207  
 DB 121 GSSPSVKCRSPKKNIOGKTLVSPQLERODSGTWCTVSOQKVEFKIDIVLAFQRA 180  
 QY 208 S 208  
 DB 181 S 181

RESULT 10  
 CD4\_ERYPA STANDARD; PRT; 397 AA.  
 ID\_008339;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Leu-3) (Fragment).  
 GN Name=CD4;  
 OS Erythrocybus patas (Red guenon) (Cercopithecus patas).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Erythrocybus.  
 NC NCBITaxID=9538;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RX MEDLINE=93049640; PubMed=1425921;  
 RA Fomsgaard A., Hirsch V.M., Johnson P.R.;  
 RT "Cloning and sequences of primate CD4 molecules: diversity of the  
 RT cellular receptor for simian immunodeficiency virus/human

RT Immunodeficiency virus.  
RL Eur. J. Immunol. 22:2973-2981(1992).  
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
CC receptor interaction. May regulate T-cell activation.  
CC -1- SUBUNIT: Associates with p56-lck (By similarity).  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

CC -----  
CC EMBL: X73324; CAAS1750.1; -.  
CC HSSP: P01730; 1WIO.  
DR GO: GO:0042101; C:T-cell receptor complex; ISS.  
DR GO: GO:0015026; F:coreceptor activity; ISS.  
DR GO: GO:0042289; F:MHC class II protein binding; ISS.  
DR GO: GO:0006955; P:immune response; ISS.  
DR GO: GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.  
DR GO: GO:0030217; P:T-cell differentiation; ISS.  
DR GO: GO:0045058; P:T-cell selection; ISS.  
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.  
DR InterPro: IPR008424; CD2.  
DR InterPro: IPR000973; CD4\_TcRg.  
DR InterPro: IPR007110; Ig-like.  
DR InterPro: IPR003596; Ig\_v.  
DR Pfam: PF00047; Ig; 2.  
DR PRINTS: PR00692; CD4TCANTIGEN.  
DR SMART: SM00406; IGV; 1.  
DR PROSITE: PS50835; IG\_LIKE; 1.  
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;  
KW Palmitate; Repeat; T-cell; Transmembrane.  
FT NON\_TER 1 369  
FT DOMAIN <1 391 Extracellular (Potential).  
FT TRANSHEM 370 391 Potential.  
FT DOMAIN 392 >397 Cytoplasmic (Potential).  
FT DOMAIN <1 98 Ig-like V-type.  
FT DOMAIN 99 176 Ig-like C2-type 1.  
FT DOMAIN 177 290 Ig-like C2-type 2.  
FT DOMAIN 291 347 Ig-like C2-type 3.  
FT CARBOHYD 269 298 N-linked (GlcNAc. . .) (By similarity).  
FT CARBOHYD 269 298 N-linked (GlcNAc. . .) (By similarity).  
FT DISULFID 14 82 By similarity.  
FT DISULFID 128 157 By similarity.  
FT DISULFID 301 343 By similarity.  
FT LIPID 392 392 S-palmitoyl cysteine (By similarity).  
FT LIPID 393 393 S-palmitoyl cysteine (By similarity).  
FT NON\_TER 397 397  
SQ SEQUENCE 397 AA; 44081 MW; 67887397A87EA4F CRC64;

Query Match 30.7%; Score 819; DB 1; Length 397;  
Best Local Similarity 87.3%; Pred. No. 5.2e-44;  
Matches 158; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 28 VVLGKGDVLTCTASQKSIQFHWKNSNQIKILGNQGSFLTKGPKSLNDRADRSRLW 87  
DB 1 VVLGKGDVLTCTASQKSIQFHWKNSNQIKILGNQGSFLTKGPKSLNDRADRSRLW 60

QY 88 DQGNFPLIINKLKIEDSDTYICEVEDQKEVQLLVFGLTANSTHLLQGSLLTLESPP 147  
DB 61 DQGNFPLIINKLKIEDSDTYICEVEDQKEVQLLVFGLTANSTHLLQGSLLTLESPP 120

QY 148 GSSPSVOCRSRPRGNKIQGKTLISVSOLELQDSGTWTCTVLONOKVEFKIDIVLAFQKA 207  
DB 121 GSSPSVOCRSRPRGNKIQGKTLISVPOLEKDSGTWTCTVSDQMTVEFKIDIVLAFQKA 180

QY 208 S 208

DB 181 S 181

RESULT 11  
ID 009261 PRELIMINARY; PRT; 397 AA.  
AC 009261;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE CD4 (Fragment).  
OS Cercopithecus sabaues.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Cercopithecus.  
NCBI\_TaxID=60711;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98017879; PubMed=9379478;  
RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,  
RA Corbet S., Barre-Sinoussi F., Allan J.S.,  
RT "Relation between phylogeny of African green monkey CD4 genes and  
RT their respective simian immunodeficiency virus genes."  
RL J. Med. Primatol. 26:120-128(1997).  
DR EMBL: AF001225; AAB60872.1; -.  
DR HSSP: P01730; 1WIO.  
DR GO: GO:0016021; C:integral to membrane; IEA.  
DR GO: GO:0007155; P:cell adhesion; IEA.  
DR GO: GO:0006955; P:immune response; IEA.  
DR InterPro: IPR000973; CD4\_TcRg.  
DR InterPro: IPR008424; CD2.  
DR InterPro: IPR007110; Ig-like.  
DR InterPro: IPR003596; Ig\_v.  
DR Pfam: PF05790; C2-sec; 2.  
DR PRINTS: PR00047; Ig; 1.  
DR PRINTS: PR00692; CD4TCANTIGEN.  
DR SMART: SM00406; IGV; 1.  
DR PROSITE: PS50835; IG\_LIKE; 1.  
FT NON\_TER 1 397  
FT NON\_TER 397 397  
SQ SEQUENCE 397 AA; 43945 MW; 80C963B92A68CD3 CRC64;

Query Match 30.1%; Score 803; DB 2; Length 397;  
Best Local Similarity 86.2%; Pred. No. 5.3e-43;  
Matches 156; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 28 VVLGKGDVLTCTASQKSIQFHWKNSNQIKILGNQGSFLTKGPKSLNDRADRSRLW 87  
DB 1 VVLGKGDVLTCTASQKSIQFHWKNSNQIKILGNQGSFLTKGPKSLNDRADRSRLW 60

QY 88 DQGNFPLIINKLKIEDSDTYICEVEDQKEVQLLVFGLTANSTHLLQGSLLTLESPP 147  
DB 61 DQGNFPLIINKLKIEDSDTYICEVEDQKEVQLLVFGLTANSTHLLQGSLLTLESPP 120

QY 148 GSSPSVOCRSRPRGNKIQGKTLISVSOLELQDSGTWTCTVLONOKVEFKIDIVLAFQKA 207  
DB 121 GSSPSVOCRSRPRGNKIQGKTLISVPOLEKDSGTWTCTVSDQMTVEFKIDIVLAFQKA 180

QY 208 S 208

DB 181 S 181

RESULT 12  
ID 009262 PRELIMINARY; PRT; 397 AA.  
AC 009262;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE CD4 (Fragment).  
NCBI\_TaxID=60711;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98017879; PubMed=9379478;  
RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,  
RA Corbet S., Barre-Sinoussi F., Allan J.S.,  
RT "Relation between phylogeny of African green monkey CD4 genes and  
RT their respective simian immunodeficiency virus genes."  
RL J. Med. Primatol. 26:120-128(1997).  
DR EMBL: AF001225; AAB60872.1; -.  
DR HSSP: P01730; 1WIO.  
DR GO: GO:0016021; C:integral to membrane; IEA.  
DR GO: GO:0007155; P:cell adhesion; IEA.  
DR GO: GO:0006955; P:immune response; IEA.  
DR InterPro: IPR000973; CD4\_TcRg.  
DR InterPro: IPR008424; CD2.  
DR InterPro: IPR007110; Ig-like.  
DR InterPro: IPR003596; Ig\_v.  
DR Pfam: PF05790; C2-sec; 2.  
DR PRINTS: PR00047; Ig; 1.  
DR PRINTS: PR00692; CD4TCANTIGEN.  
DR SMART: SM00406; IGV; 1.  
DR PROSITE: PS50835; IG\_LIKE; 1.  
FT NON\_TER 1 397  
FT NON\_TER 397 397  
SQ SEQUENCE 397 AA; 43945 MW; 80C963B92A68CD3 CRC64;

OS Cercopithecus tantalus.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; Cercopithecus.  
 OX NCBI\_TaxId=60712;  
 RN [1]  
 RP MEDLINE=98017879; PubMed=9379478;  
 RX Fomesgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,  
 RA Corbet S., Barre-Sinoussi F., Allan J.S.;  
 RT "Relation between phylogeny of African green monkey CD4 genes and  
 their respective simian immunodeficiency virus genes.";  
 RL J. Med. Primatol. 26:120-128(1997).  
 DR EMBL; AF001221; AAB60868.1; -.  
 DR HSSP; P01730; 1WIO.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR GO; GO:0006955; P:immune response; IEA.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TcAg.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003596; Ig v.  
 DR Pfam; PF005790; C2-sec; 2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IgV; 1.  
 DR PROSITE; PSS0835; IG\_LIKE; 1.  
 FT NON\_TER 1 397  
 FT NON\_TER 1 397  
 SQ SEQUENCE 397 AA; 43954 MW; CF7F2F5D82335BD0 CRC64;

Query Match 30.1%; Score 803; DB 2; Length 397;  
 Best Local Similarity 86.2%; Pred. No. 5.3e-43;  
 Matches 156; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 28 VVLGKKDPTVELTCTASOKKSIOFHWNKSNQIKILGQGSFLTKGSSKLRDRIDRSRLW 87  
 DB 1 VVLGKKDPTVELTCTASOKKSIOFHWNKSNQIKILGQGSFLTKGSSKLRDRIDRSRLW 60  
 QY 88 DQGNFPLIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANSPTHLQGSQSLTLLTLESP 147  
 DB 61 DQGNFPLIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANSPTHLQGSQSLTLLTLESP 120  
 QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWTCTVLAQOKVEFKIDIVLAFOKA 207  
 DB 121 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWTCTVLAQOKVEFKIDIVLAFOKA 180  
 QY 208 S 208  
 DB 181 S 181  
 RESULT 13  
 009263 PRELIMINARY; PRT; 397 AA.  
 AC 009263;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE CD4 (Fragment).  
 GN Name=CD4;  
 OS Cercopithecus tantalus.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; Cercopithecus.  
 OX NCBI\_TaxId=60712;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98017879; PubMed=9379478;  
 RA Fomesgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,  
 RA Corbet S., Barre-Sinoussi F., Allan J.S.;  
 RT "Relation between phylogeny of African green monkey CD4 genes and  
 their respective simian immunodeficiency virus genes.";

RL J. Med. Primatol. 26:120-128(1997).  
 DR EMBL; AF001222; AAB60869.1; -.  
 DR HSSP; P01730; 1WIO.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR GO; GO:0006955; P:immune response; IEA.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TcAg.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003596; Ig v.  
 DR Pfam; PF005790; C2-sec; 2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IgV; 1.  
 DR PROSITE; PSS0835; IG\_LIKE; 1.  
 FT NON\_TER 1 397  
 FT NON\_TER 1 397  
 SQ SEQUENCE 397 AA; 43994 MW; A3CD031535A51524 CRC64;

Query Match 30.1%; Score 803; DB 2; Length 397;  
 Best Local Similarity 86.2%; Pred. No. 5.3e-43;  
 Matches 156; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 28 VVLGKKDPTVELTCTASOKKSIOFHWNKSNQIKILGQGSFLTKGSSKLRDRIDRSRLW 87  
 DB 1 VVLGKKDPTVELTCTASOKKSIOFHWNKSNQIKILGQGSFLTKGSSKLRDRIDRSRLW 60  
 QY 88 DQGNFPLIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANSPTHLQGSQSLTLLTLESP 147  
 DB 61 DQGNFPLIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANSPTHLQGSQSLTLLTLESP 120  
 QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWTCTVLAQOKVEFKIDIVLAFOKA 207  
 DB 121 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWTCTVLAQOKVEFKIDIVLAFOKA 180  
 QY 208 S 208  
 DB 181 S 181  
 RESULT 14  
 095NE9 PRELIMINARY; PRT; 397 AA.  
 AC 095NE9;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE CD4 (Fragment).  
 GN Name=CD4;  
 OS Cercopithecus pygerythrus (Vervet monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; Cercopithecus.  
 OX NCBI\_TaxId=60710;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98017879; PubMed=9379478;  
 RA Fomesgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,  
 RA Corbet S., Barre-Sinoussi F., Allan J.S.;  
 RT "Relation between phylogeny of African green monkey CD4 genes and  
 their respective simian immunodeficiency virus genes.";  
 RL J. Med. Primatol. 26:120-128(1997).  
 DR EMBL; AF001227; AAB60874.1; -.  
 DR HSSP; P01730; 1WIO.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR GO; GO:0006955; P:immune response; IEA.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TcAg.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003596; Ig v.  
 DR Pfam; PF005790; C2-sec; 2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.

DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PSS0835; IG-LIKE; 1.  
 FT NON\_TER 1  
 FT SEQUENCE 397 AA; 43946 MW; 21C3B30882ABFBC0 CRC64;

Query Match 30.1%; Score 803; DB 2; Length 397;  
 Best Local Similarity 86.2%; Pred. No. 5.3e-43;  
 Matches 156; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 28 VVLGKGDVLTCTASQKSIOPHWKNSNOIKILNQGSLFYTGPSKLANDRADSRSLW 87  
 DB 1 VVLGKGDVLTCTASQKSIOPHWKNSNOIKILNQGSLFYTGPSKLANDRADSRSLW 60

QY 88 DQGNPPLITKNLKIETSDTYICEVDQKEVQLLVFGLTANSPDTHLLOQOSLTLTLESPP 147  
 DB 61 DQGNPPLITKNLKIETSDTYICEVDQKEVQLLVFGLTANSPDTHLLOQOSLTLTLESPP 120

QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLQONQKVEFKIDIVVLAFOKA 207  
 DB 121 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLQONQKVEFKIDIVVLAFOKA 180

QY 208 S 208  
 DB 181 S 181

RESULT 15

ID 009259 PRELIMINARY; PRT; 397 AA.

AC 009259;  
 DT 01-JUL-1997 (TReMBLrel. 04, Created)

DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

DE CD4 (Fragment).

GN Name=CD4;

OS Cercopithecus sabaus.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;

OC Cercopithecinae; Cercopithecus.

OX NCBI\_TaxID=60711;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98017879; PubMed=9379478;

RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,

RA Corbet S., Barre-Sinoussi F., Allan J.S.;

RT "Relation between phylogeny of African green monkey CD4 genes and

RT their respective simian immunodeficiency virus genes.";

RL J. Med. Primatol. 26:120-128(1997).

DR EMBL; AF001223; AAB60870.1; --

DR HSBP; P01730; IWIQ.

DR GO; GO:0016021; C:Integral to membrane; IEA.

DR GO; GO:0007155; P:cell adhesion; IEA.

DR GO; GO:0006955; P:immune response; IEA.

DR InterPro; IPR008424; CD2

DR InterPro; IPR00973; CD4\_TcAg.

DR InterPro; IPR007110; Ig-Like.

DR Pfam; PF05790; C2-set; 2.

DR Pfam; PF00047; Ig; 1.

DR PRINTS; PR00692; CD4TCANTIGEN.

DR SMART; SM00406; IGV; 1.

DR PROSITE; PSS0835; IG-LIKE; 1.

FT NON\_TER 1

FT SEQUENCE 397 AA; 43881 MW; 7CB39AD0F8506C81 CRC64;

QY 28 VVLGKGDVLTCTASQKSIOPHWKNSNOIKILNQGSLFYTGPSKLANDRADSRSLW 87

Query Match 29.9%; Score 799; DB 2; Length 397;

Best Local Similarity 85.6%; Pred. No. 9.6e-43;

Matches 155; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

QY 28 VVLGKGDVLTCTASQKSIOPHWKNSNOIKILNQGSLFYTGPSKLANDRADSRSLW 87

DB 1 VVLGKGDVLTCTASQKSIOPHWKNSNOIKILNQGSLFYTGPSKLANDRADSRSLW 60  
 QY 88 DQGNPPLITKNLKIETSDTYICEVDQKEVQLLVFGLTANSPDTHLLOQOSLTLTLESPP 147  
 DB 61 DQGNPPLITKNLKIETSDTYICEVDQKEVQLLVFGLTANSPDTHLLOQOSLTLTLESPP 120  
 QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLQONQKVEFKIDIVVLAFOKA 207  
 DB 121 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLQONQKVEFKIDIVVLAFOKA 180  
 QY 208 S 208  
 DB 181 S 181

Search completed: September 20, 2005, 16:09:30  
 Job time : 141 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2005, 08:29:41 ; Search time 6457 Seconds  
(without alignments)  
10806.181 Million cell updates/sec

Title: US-09-936-702-4

Perfect score: 1440  
Sequence: 1 atgaaccgsg9gagtcctt.....gtccagcagtaacatactgg 1440

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

GenBml:.\*  
1: gb\_ba:.\*  
2: gb\_hcg:.\*  
3: gb\_in:.\*  
4: gb\_om:.\*  
5: gb\_ov:.\*  
6: gb\_pat:.\*  
7: gb\_ph:.\*  
8: gb\_pl:.\*  
9: gb\_pr:.\*  
10: gb\_ro:.\*  
11: gb\_rts:.\*  
12: gb\_sy:.\*  
13: gb\_un:.\*  
14: gb\_vl:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1440	100.0	1440	6	BD268881 Novel chl
2	625	43.4	1213	6	AR363716 Sequence
3	624.4	43.4	1377	6	AX100880 Sequence
4	624.4	43.4	1742	6	AR380468 Sequence
5	624.4	43.4	1742	6	AX287109 Sequence
6	624.4	43.4	1742	9	HUMATCTA
7	624.4	43.4	1910	6	HUMATCTA
8	624.4	43.4	1910	6	108116 Sequence 3
9	624.4	43.4	3133	6	108115 Sequence 1
10	624.4	43.4	7533	12	AY438650 Cloning v
11	622.8	43.2	1273	6	106223 Sequence 1
12	622.8	43.2	1273	6	107147 Sequence 1
13	622.8	43.2	1304	6	AR067943 Sequence
14	622.8	43.2	1304	6	AR166802 Sequence
15	622.8	43.2	1304	6	AR560088 Sequence
16	622.8	43.2	1389	6	AR062468 Sequence
17	622.8	43.2	1389	6	AR067924 Sequence
18	622.8	43.2	1389	6	AR166783 Sequence
19	622.8	43.2	1389	6	AR560069 Sequence

20	622.8	43.2	1537	6	CQ722428 Sequence
21	622.8	43.2	1539	6	AR062469 Sequence
22	622.8	43.2	1599	6	AR067925 Sequence
23	622.8	43.2	1599	6	AR166784 Sequence
24	622.8	43.2	1599	6	AR560070 Sequence
25	622.8	43.2	1728	6	AR062467 Sequence
26	622.8	43.2	1728	6	AR067923 Sequence
27	622.8	43.2	1728	6	AR166782 Sequence
28	622.8	43.2	1728	6	AR560068 Sequence
29	622.8	43.2	1742	6	AR104143 Sequence
30	622.8	43.2	1742	6	108673 Sequence 2
31	622.8	43.2	1742	6	AR339862 Sequence
32	622.8	43.2	1742	6	AR448985 Sequence
33	622.8	43.2	2589	6	AR392157 Sequence
34	622.8	43.2	3084	9	BC025782 Homo sapi
35	621.2	43.1	1273	6	AR104142 Sequence
36	621.2	43.1	1273	6	AR448979 Sequence
37	621.2	43.1	1416	6	107030 Sequence 11
38	621.2	43.1	1416	6	109301 Sequence 1
39	621.2	43.1	1416	6	127613 Sequence 2
40	621.2	43.1	1416	6	AR489592 Sequence
41	621.2	43.1	1742	6	107208 Sequence 4
42	621.2	43.1	2940	6	AR064430 Sequence
43	621.2	43.1	3064	6	109237 Sequence 1
44	621.2	43.1	3426	6	AR064429 Sequence
45	621.2	43.1	3480	6	AR064431 Sequence

#### ALIGNMENTS

RESULT 1  
BD268881 1440 bp DNA linear PAT 17-JUL-2003  
LOCUS Novel chimeric protein for prevention and treatment of HIV  
DEFINITION

BD268881  
ACCESSION BD268881.1 GI:33078649  
VERSION JP 2002538814-A/1.  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 1440)  
Berger,E.A. and Castillo,C.D.  
TITLE Novel chimeric protein for prevention and treatment of HIV

#### JOURNAL

Patent: JP 2002538814-A 1 19-NOV-2002;  
THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY  
THE AUSTRIAN NORDIC BIOTHERAPEUTICS AKTIENGESELLSCHAFT SECRETARY  
DEPARTMENT OF HEALTH AND HUMAN SERVICES THE NATIONAL INSTITUTES OF  
HEALTH

#### COMMENT

OS Artificial Sequence  
PN JP 2002538814-A/1  
PD 19-NOV-2002  
PF 16-MAR-2000 JP 2000605633  
PI EDWARD A. BERGER, CHRISTIE DEL CASTILLO  
PC C12N15/09,A61K38/00,A61P31/18,C07K5/103,C07K14/00,C07K14/155,  
PC C07K14/73,C07K16/10,C07K19/00,C12N5/10,C12P21/02// (C12P21/02,  
PC C12N15/00,C12N5/00,A61K37/02  
PC C12N15/00,C12N5/00,A61K37/02  
CC Description of Artificial Sequence: CDA-seFv(17b) FH Key  
Location/Qualifiers  
FT source 1..1440  
FT Location/Qualifiers  
1..1440  
/organism='Artificial Sequence'.  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

#### FEATURES

source

#### ORIGIN

Query Match 100.0%; Score 1440; DB 6; Length 1440;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGACCGGGAGGCCCTTTTAGGCACTTGCTGTGTGTGCAACCTGGGCGCTCCCTCCA
DB 1 ATGACCGGGAGGCCCTTTTAGGCACTTGCTGTGTGTGCAACCTGGGCGCTCCCTCCA
QY 61 GCAGCCACTCAGGAGAAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC
DB 61 GCAGCCACTCAGGAGAAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC
QY 121 TGTACAGCTTCCAGAAAGAGCATTCATTCCTGAAAACCTCAACCATTAAG
DB 121 TGTACAGCTTCCAGAAAGAGCATTCATTCCTGAAAACCTCAACCATTAAG
QY 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGGTGCATCCAAAGCTGAATGATCGGCT
DB 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGGTGCATCCAAAGCTGAATGATCGGCT
QY 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG
DB 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG
QY 301 ATGAGAGACTCAAGTACTTACATCTGTGAAGTGAAGACCAAGAAAGAGTGCATTTG
DB 301 ATGAGAGACTCAAGTACTTACATCTGTGAAGTGAAGACCAAGAAAGAGTGCATTTG
QY 361 CTAGTGTGGGATTTAGCTGCAACTGTGACACCCAGCTGCTAGGGGCGAGGCTGACC
DB 361 CTAGTGTGGGATTTAGCTGCAACTGTGACACCCAGCTGCTAGGGGCGAGGCTGACC
QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGATCCAAAGGCT
DB 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGATCCAAAGGCT
QY 481 AAAAACCTACAGGGGGGAAAGACCTCTCGGTGTCTAGCTGGAAGCTCCAGATAGTGGC
DB 481 AAAAACCTACAGGGGGGAAAGACCTCTCGGTGTCTAGCTGGAAGCTCCAGATAGTGGC
QY 541 ACCGAGCATGCACTGTCTTGCAAGACCAAGAAAGGTGAGATTCAAAATAGACATGTTG
DB 541 ACCGAGCATGCACTGTCTTGCAAGACCAAGAAAGGTGAGATTCAAAATAGACATGTTG
QY 601 GTGTAGCTTTCCAGAAAGGCTCCGAGGTGCGCGTGTGAGTGGGGAGCGGTTCAAGCGGA
DB 601 GTGTAGCTTTCCAGAAAGGCTCCGAGGTGCGCGTGTGAGTGGGGAGCGGTTCAAGCGGA
QY 661 GTGTGATCCGGTGGCGGAGGTCGGGCGGGGTGGAACCGGGGTGGCGGCTCCGAGGCT
DB 661 GTGTGATCCGGTGGCGGAGGTCGGGCGGGGTGGAACCGGGGTGGCGGCTCCGAGGCT
QY 721 GAGAGTTCAGAGGTGCAAGTGTGAGTGTGGGGCTAGAGTGAAGAACTCGGGTCTCTCG
DB 721 GAGAGTTCAGAGGTGCAAGTGTGAGTGTGGGGCTAGAGTGAAGAACTCGGGTCTCTCG
QY 781 GTGAGGTCTCTGCAAGGCTCTGAGAGACCTTCATCAGATTAATTTTAACTCTGGTG
DB 781 GTGAGGTCTCTGCAAGGCTCTGAGAGACCTTCATCAGATTAATTTTAACTCTGGTG
QY 841 GTGAGGTCTCTGCAAGGCTCTGAGAGACCTTCATCAGATTAATTTTAACTCTGGTG
DB 841 GTGAGGTCTCTGCAAGGCTCTGAGAGACCTTCATCAGATTAATTTTAACTCTGGTG
QY 901 GCACTAGGCAACCGGCTCTCAAGGAGAGTCAAGTTACCGGGGCAAGTCCACAGAGC
DB 901 GCACTAGGCAACCGGCTCTCAAGGAGAGTCAAGTTACCGGGGCAAGTCCACAGAGC
QY 961 ACAGTCTACCTGAGCTGCGGAATCTAAGATCTGACGATACGCGCGTATATTTCTGTGCG
DB 961 ACAGTCTACCTGAGCTGCGGAATCTAAGATCTGACGATACGCGCGTATATTTCTGTGCG
QY 1021 GAGGTGTACGAGGAGGCGGAGCGAGGCGGAATATGATTAATTAATGGGTTTCTGAACAT
DB 1021 GAGGTGTACGAGGAGGCGGAGCGAGGCGGAATATGATTAATTAATGGGTTTCTGAACAT

```

RESULT 2  
 AR363716 1213 bp DNA linear PAT 03-SEP-2003  
 LOCUS AR363716 Sequence 1 from patent US 5223418.  
 DEFINITION AR363716  
 VERSION AR363716.1 GI:34425655  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1213)  
 AUTHORS Arcuri,E.J., Brawner,M.E., Donovan,M.J., Gerber,R.G. and Keller,J.A.  
 TITLE Method of improving the yield of heterologous proteins produced by Streptomyces lividans  
 JOURNAL Patent: US 5223418-A 1 29-JUN-1993;  
 FEATURES Location/Qualifiers  
 1..1213  
 /organism="unknown"  
 /mol\_type="genomic DNA"

ORIGIN

Query Match 43.4%; Score 625; DB 6; Length 1213;  
 Best Local Similarity 97.7%; Pred. No. 1,3e-134;  
 Matches 634; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

QY 1 ATGACCGGGAGTCCCTTTTAGGCACTTGCTGTGTGTGCAACTGGCGCTCTCCCA
DB 79 ATGACCGGGAGTCCCTTTTAGGCACTTGCTGTGTGTGCAACTGGCGCTCTCCCA
QY 61 GCAGCCACTCAGGAGAAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC
DB 61 GCAGCCACTCAGGAGAAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC
QY 121 TGTACAGCTTCCAGAAAGAGCATTCATTCCTGAAAACCTCAACCATTAAG
DB 121 TGTACAGCTTCCAGAAAGAGCATTCATTCCTGAAAACCTCAACCATTAAG
QY 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTGCATCCAAAGCTGAATGATCGGCT
DB 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTGCATCCAAAGCTGAATGATCGGCT
QY 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG
DB 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG
QY 319 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG
DB 319 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG

```

QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGACCAAGAGAGAGTGTCAATTG 360  
 Db 379 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGACCAAGAGAGAGTGTCAATTG 438  
 QY 361 CTAGTGTGGATTGACTGCGCACTGTGACACCACTGTCTTCAAGGGGCAAGGCTGACC 420  
 Db 439 CTAGTGTGGATTGACTGCGCACTGTGACACCACTGTCTTCAAGGGGCAAGGCTGACC 498  
 QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCATGTGCAATTTAGAGATCCAAAGGGGT 480  
 Db 499 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCATGTGCAATTTAGAGATCCAAAGGGGT 558  
 QY 481 AAAAACAATACAGGGGGGAGAACCCCTCTCGGTGTCTGAGCTGAGAGCTCCAGATATGAGGC 540  
 Db 559 AAAAACAATACAGGGGGGAGAACCCCTCTCGGTGTCTGAGCTGAGAGCTCCAGATATGAGGC 618  
 QY 541 ACCTGACATGACACTGTCTTGACAGAACCAAGAGAGTGTGCAAAATATGACATCGTG 600  
 Db 619 ACCTGACATGACACTGTCTTGACAGAACCAAGAGAGTGTGCAAAATATGACATCGTG 678  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGAGAGTGGCGGTAGTGGGAGGCG 649  
 Db 679 GTGCTAGCTTTCCAGAAAGGCTCCGAGAGTGTGTAAGAAAGAGGCG 727

RESULT 3  
 AX100880  
 LOCUS AX100880 1377 bp DNA linear PAT 10-APR-2001  
 DEFINITION Sequence 30 from Patent WO0122084.  
 ACCESSION AX100880  
 VERSION AX100880.1 GI:13619792

KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Jakobsen, B. K.  
 TITLE SpR identification of inhibitors of receptor-ligand interactions  
 JOURNAL Patent: WO 0122084-A 30 29-MAR-2001;  
 Avidex Ltd (GB)

FEATURES  
 source location/Qualifiers  
 1..1377  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 1..1377  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC36495.1"  
 /db\_xref="GI:13619793"

CDS

ORIGIN  
 Query Match 43.4%; Score 624.4; DB 6; Length 1377;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-134;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTGTGGTGTGCAACTGGGCGCTCTCCCA 60  
 Db 1 ATGAACCGGGAGTCCCTTTAGGCACTTGTGGTGTGCAACTGGGCGCTCTCCCA 60  
 QY 61 GCAGCCACTCAGGAGAAAGAGTGTGTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
 Db 61 GCAGCCACTCAGGAGAAAGAGTGTGTGGGCAAAAAGGGGATACAGTGAACCTGACC 120

Db 61 GCAGCCACTCAGGAGAAAGAGTGTGTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
 QY 121 TGTACAGCTTCCAGAGAAAGAGCATATCAATTCCATGGAATAAACTCCAAACGATTAAG 180  
 Db 121 TGTACAGCTTCCAGAGAAAGAGCATATCAATTCCATGGAATAAACTCCAAACGATTAAG 180  
 QY 181 ATTCTGGAAATCAGGGCTCTCTTTAACTAAAGTTCATCCAAAGTGAATGATCGGCT 240  
 Db 181 ATTCTGGAAATCAGGGCTCTCTTTAACTAAAGTTCATCCAAAGTGAATGATCGGCT 240  
 QY 241 GACTCAAGAAAGAGCTTTGGAGCAAGGAAATTTCCCTGATCATCAAGAAATCTTAAG 300  
 Db 241 GACTCAAGAAAGAGCTTTGGAGCAAGGAAATTTCCCTGATCATCAAGAAATCTTAAG 300  
 QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGACCAAGAGAGAGTGTCAATTG 360  
 Db 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGACCAAGAGAGAGTGTCAATTG 360  
 QY 361 CTAGTGTGGATTGACTGCGCACTGTGACACCACTGTCTTCAAGGGGCAAGGCTGACC 420  
 Db 361 CTAGTGTGGATTGACTGCGCACTGTGACACCACTGTCTTCAAGGGGCAAGGCTGACC 420  
 QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCATGTGCAATTTAGAGATCCAAAGGGGT 480  
 Db 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCATGTGCAATTTAGAGATCCAAAGGGGT 480  
 QY 481 AAAAACAATACAGGGGGGAGAACCCCTCTCGGTGTCTGAGCTGAGAGCTCCAGATATGAGGC 540  
 Db 481 AAAAACAATACAGGGGGGAGAACCCCTCTCGGTGTCTGAGCTGAGAGCTCCAGATATGAGGC 540  
 QY 541 ACCTGACATGACACTGTCTTGACAGAACCAAGAGAGTGTGCAAAATATGACATCGTG 600  
 Db 541 ACCTGACATGACACTGTCTTGACAGAACCAAGAGAGTGTGCAAAATATGACATCGTG 600  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGAGAGTGGCGGTAGTGGGAGGCG 649  
 Db 601 GTGCTAGCTTTCCAGAAAGGCTCCGAGAGTGTGTAAGAAAGAGGCG 626

RESULT 4  
 AR380468  
 LOCUS AR380468 1742 bp DNA linear PAT 18-DEC-2003  
 DEFINITION Sequence 1013 from patent US 6607879.  
 ACCESSION AR380468  
 VERSION AR380468.1 GI:40088102  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1742)  
 AUTHORS Cocke, B. G., Stuart, S. G. and Seilhamer, J. J.  
 TITLE Compositions for the detection of blood cell and immunological  
 response gene expression  
 JOURNAL Patent: US 6607879-A 1013 19-AUG-2003;  
 FEATURES location/Qualifiers  
 1..1742  
 /organism="unknown"  
 /mol\_type="genomic DNA"

ORIGIN

Query Match 43.4%; Score 624.4; DB 6; Length 1742;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-134;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTGTGGTGTGCAACTGGGCGCTCTCCCA 60  
 Db 76 ATGAACCGGGAGTCCCTTTAGGCACTTGTGGTGTGCAACTGGGCGCTCTCCCA 135  
 QY 61 GCAGCCACTCAGGAGAAAGAGTGTGTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
 Db 136 GCAGCCACTCAGGAGAAAGAGTGTGTGGGCAAAAAGGGGATACAGTGAACCTGACC 195  
 QY 121 TGTACAGCTTCCAGAGAAAGAGCATATCAATTCCATGGAATAAACTCCAAACGATTAAG 180

```
Db      196 TGTACAGCTTCCAGAAAGAGATACATTCCACTCGAAGAAAACTCCAAACGATTAAG 255
Qy      181 ATTCTGGAAATCAGGGCTCTCTTTAACTAAAGGTCATCAAGCTGAATGATCGGCT 240
Db      256 ATTCGGGAATCAGGGCTCTCTTTAACTAAAGGTCATCAAGCTGAATGATCGGCT 315
Qy      241 GACTCAAGAAAGACCTTTGGGACCAAGAAACCTTCCCTGATCATCAAGAACTTTAAG 300
Db      316 GACTCAAGAAAGACCTTTGGGACCAAGAAACCTTCCCTGATCATCAAGAACTTTAAG 375
Qy      301 ATGAAGACTCAATCTTAATCTTGAAGTGAAGGACCAAGAAAGGAGGTGCATTG 360
Db      376 ATGAAGACTCAATCTTAATCTTGAAGTGAAGGACCAAGAAAGGAGGTGCATTG 435
Qy      361 CTAGTGTGGATGACTGCACTGCAACCTGTCAGACCCAGCTTCAGGGGACAGCTGACC 420
Db      436 CTAGTGTGGATGACTGCACTGCAACCTGTCAGACCCAGCTTCAGGGGACAGCTGACC 495
Qy      421 CTGACCTTGGAGAGCCCCCTGGTAGTACCCCTCACTGCAATGTAGAGTCCAAAGGGT 480
Db      496 CTGACCTTGGAGAGCCCCCTGGTAGTACCCCTCACTGCAATGTAGAGTCCAAAGGGT 555
Qy      481 AAAAATATCAGGGGGGGAAGACCCCTCTCGTGTCTCACTGAGCTCCAGATAGTGGC 540
Db      556 AAAAATATCAGGGGGGGAAGACCCCTCTCGTGTCTCACTGAGCTCCAGATAGTGGC 615
Qy      541 ACCTGACATGCACTGTCTTGGCAAGCAAGAAAGGTGAGTTCAAAATAGACATCGTG 600
Db      616 ACCTGACATGCACTGTCTTGGCAAGCAAGAAAGGTGAGTTCAAAATAGACATCGTG 675
Qy      601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
Db      676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701

RESULT 5
AX287109      1742 bp      DNA      linear      PAT 21-NOV-2001
LOCUS      AX287109
DEFINITION      Sequence 7 from Patent WO0164752.
ACCESSION      AX287109
VERSION      AX287109.1 GI:17049085
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS      Litchman,D.R., Kwon,D., van Kooyk,Y.C. and Geijtenbeek,T.C.
TITLE      Antibody inhibiting the binding between gp120 and dc-sigin and
JOURNAL      screening methods
PATENT      Patent: WO 0164752-A 7 07-SEP-2001;
FEATURES      location/Qualifiers
source      1..1742
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      43.4%; Score 624.4; DB 6; Length 1742;
Best Local Similarity 99.8%; Pred. No. 1.8e-134;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ATGAACCGGGAGTCCCTTTAGCACTTGTGTGTCGAACCTGGGCTCTCCCA 60
Db      76 ATGAACCGGGAGTCCCTTTAGCACTTGTGTGTCGAACCTGGGCTCTCCCA 135
Qy      61 GCAGCCACTCAGGAAAGAAAGTGTCTGGGCAAAAAAGGGATACAGTGAACCTGACC 120
Db      136 GCAGCCACTCAGGAAAGAAAGTGTCTGGGCAAAAAAGGGATACAGTGAACCTGACC 155
Qy      121 TGTACAGCTTCCAGAAAGAGCATATTCACCTGCAAGAAAACTCCAAACGATTAAG 180
```

```
Db      196 TGTACAGCTTCCAGAAAGAGCATACATTCCATGAAAAAACTCCAAACGATTAAG 255
Qy      181 ATTCTGGAAATCAGGGCTCTCTTTAACTAAAGGTCATCCAAAGCTGAATGATCGGCT 240
Db      256 ATTCGGGAATCAGGGCTCTCTTTAACTAAAGGTCATCCAAAGCTGAATGATCGGCT 315
Qy      241 GACTCAAGAAAGACCTTTGGGACCAAGAAACCTTCCCTGATCATCAAGAACTTTAAG 300
Db      316 GACTCAAGAAAGACCTTTGGGACCAAGAAACCTTCCCTGATCATCAAGAACTTTAAG 375
Qy      301 ATGAAGACTCAATCTTAATCTTGAAGTGAAGGACCAAGAAAGGAGGTGCATTG 360
Db      376 ATGAAGACTCAATCTTAATCTTGAAGTGAAGGACCAAGAAAGGAGGTGCATTG 435
Qy      361 CTAGTGTGGATGACTGCACTGCAACCTGTCAGACCCAGCTTCAGGGGACAGCTGACC 420
Db      436 CTAGTGTGGATGACTGCACTGCAACCTGTCAGACCCAGCTTCAGGGGACAGCTGACC 495
Qy      421 CTGACCTTGGAGAGCCCCCTGGTAGTACCCCTCACTGCAATGTAGAGTCCAAAGGGT 480
Db      496 CTGACCTTGGAGAGCCCCCTGGTAGTACCCCTCACTGCAATGTAGAGTCCAAAGGGT 555
Qy      481 AAAAATATCAGGGGGGGAAGACCCCTCTCGTGTCTCACTGAGCTCCAGATAGTGGC 540
Db      556 AAAAATATCAGGGGGGGAAGACCCCTCTCGTGTCTCACTGAGCTCCAGATAGTGGC 615
Qy      541 ACCTGACATGCACTGTCTTGGCAAGCAAGAAAGGTGAGTTCAAAATAGACATCGTG 600
Db      616 ACCTGACATGCACTGTCTTGGCAAGCAAGAAAGGTGAGTTCAAAATAGACATCGTG 675
Qy      601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
Db      676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701

RESULT 6
HUMATCT4      1742 bp      mRNA      linear      PRI 27-APR-1993
LOCUS      HUMATCT4
DEFINITION      Human T-cell surface glycoprotein T4 mRNA, complete cds.
ACCESSION      M12807
VERSION      M12807.1 GI:179141
KEYWORDS      glycoprotein; immunoglobulin super gene family; recognition
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS      Maddon,P.J., Litchman,D.R., Godfrey,M., Maddon,D.E., Chess,L. and
TITLE      Axel,R.
JOURNAL      The isolation and nucleotide sequence of a cDNA encoding the T cell
MEDLINE      surface protein T4; a new member of the immunoglobulin gene family
PUBMED      Cell 42 (1), 93-104 (1985)
85254948
2990730
2 (bases 153 to 153)
REFERENCE      Litchman,D.R.
AUTHORS      Unpublished (1986)
JOURNAL
COMMENT      Original source text: Human peripheral T lymphocyte, cDNA to mRNA,
clone pT4B.
T4 is a member of the immunoglobulin supergene family. The T4 mRNA
encodes contiguous V- and J-like elements without the requirement
for DNA recombination events. The V-like domain can be found at
positions 145 to 426 and the J-like domain from 427 to 471. The T4
protein may serve as the specific surface receptor for the AIDS
virus [1].
With the revision of the nucleotide at position 153, the amino acid
encoded by 151-153 becomes lysine. This is now regarded to be the
first residue in the mature protein [2].
location/Qualifiers
1..1742
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/organism="Homo sapiens"
```

CDS

/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
76..1452  
/note="T4 surface glycoprotein precursor"  
/codon\_start=1  
/protein\_id="AAA3572.1"  
/db\_xref="GI:179142"  
/translation="MNRGVPRHLLVLQALLPAATQGRKVLGKKDVTVEITCTAS  
QKSIQPHMKNNOIKILGNQSFITGSPKLNDRADSRSLMDQNPPLIKNLKIE  
DSPTVCEVEDQKEEVLVFLGTANSDFLILQOSLTLSPSSPSVOCRRSPRG  
KNIOGGKTLVSQLELDGSMCTVIONOKVEFKIDIVLAFORASSIVYKKEGEO  
VERSPPLATFVERKLTGSGELMOWAERASSKSMITPDLKNKESYKRVTDPKIQMK  
KLPHILTPQALPQYAGSNLTLALEKTKLHQEVNLVVRATQLOKNTLCEWGT  
SPKLMISLKLKNEKAVSKREKAVWVNLNPAQMWOCILSDSQVLLSNIKVLPTWST  
PVQPMALIVLGVALILFLIGLIFFCVCRHRRROAERMSQIKRLSEKKTCCPBR  
FOKTCSPPI"  
76..150  
/note="T4 surface glycoprotein signal peptide"  
151..1449  
/product="T4 surface glycoprotein mature peptide"

sig\_peptide  
mat\_peptide  
198 bp upstream of Real site.

Query Match 43.4%; Score 624.4; DB 9; Length 1742;  
Best Local Similarity 99.8%; Pred. No. 1.8e-134;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAAACCGGGAGAGTCCCTTTAGGCACTTGTGGTGTGCAACTGGCGCTCTCCCA 60  
76 ATGAACCGGGAGAGTCCCTTTAGGCACTTGTGGTGTGCAACTGGCGCTCTCCCA 135  
61 GCAGCCACTCAGGAAAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAACCTGACC 120  
136 GCAGCCACTCAGGAAAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAACCTGACC 195  
121 TGTACAGCTTCCAGAAAGAGATCAATTCACAGTGAACCACTCCAGATTAAG 180  
196 TGTACAGCTTCCAGAAAGAGATCAATTCACAGTGAACCACTCCAGATTAAG 255  
181 ATTCTGGGAATCAGGGCTCTTCTTAATAAGAGTCCATCAAGCTGAATGATCGGCT 240  
256 ATTCTGGGAATCAGGGCTCTTCTTAATAAGAGTCCATCAAGCTGAATGATCGGCT 315  
241 GACTCAAGAAAGCCTTTGGGACCAAGAAATTTCCCTGATCAAGAAATCTTAAG 300  
316 GACTCAAGAAAGCCTTTGGGACCAAGAAATTTCCCTGATCAAGAAATCTTAAG 375  
301 ATAGAAGACTCAATATCTTATCTGTGAAGTGAAGACCAAGAAAGAGTGTCAATTG 360  
376 ATAGAAGACTCAATATCTTATCTGTGAAGTGAAGACCAAGAAAGAGTGTCAATTG 435  
361 CTAGTGTTCGATTTAGTGCACCTGACACCACTGCTTACAGGGGAGAGCTGACC 420  
436 CTAGTGTTCGATTTAGTGCACCTGACACCACTGCTTACAGGGGAGAGCTGACC 495  
421 CTAGCTTGGAGAGCCCCCTGTGATAGCCCCCTCACTGCAATGATAGAGTCAAGGGT 480  
496 CTAGCTTGGAGAGCCCCCTGTGATAGCCCCCTCACTGCAATGATAGAGTCAAGGGT 555  
481 AAAAATCAATAGGGGGGAGAACCTCTCGTGTCTCACTGCAATGATAGAGTCAAGGGT 540  
556 AAAAATCAATAGGGGGGAGAACCTCTCGTGTCTCACTGCAATGATAGAGTCAAGGGT 615  
541 ACCTGAGATGACCTGTCTTGAAGACCAAGAAAGTGAAGTTCAAAATAGACATCTGTG 600  
616 ACCTGAGATGACCTGTCTTGAAGACCAAGAAAGTGAAGTTCAAAATAGACATCTGTG 675  
601 GTCTAGCTTTCCAGAAAGGCTCCGG 626  
676 GTCTAGCTTTCCAGAAAGGCTCCAG 701

RESULT 7  
HUMATCT4A

LOCUS HUMATCT4A 1742 bp DNA linear PRI 24-JAN-1994  
DEFINITION Human T4 surface glycoprotein CD4 gene, complete cds.  
ACCESSION M35160  
VERSION M35160.1 GI:179143  
KEYWORDS glycoprotein; immunoglobulin super gene family; recognition antigen; surface antigen.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1742)  
AUTHORS Hodge,T.W., Sasso,D.R. and McDougal,J.S.  
TITLE Humans with OKT4-epitope deficiency have a single nucleotide base change in the CD4 gene, resulting in substitution of TRP240 for ARG240  
JOURNAL Hum. Immunol. 30 (2), 99-104 (1991)  
MEDLINE 91216786  
FUBMED 1708753  
REFERENCE 2 (bases 1 to 1742)  
AUTHORS Hodge,T.W.  
TITLE Direct Submission  
JOURNAL Submitted (12-JUN-1990) T.W. Hodge, A-25 Bldg. 1, Rm. 1226, Centers for Disease Control, 1600 Clifton Rd. Atlanta, GA 30333 USA  
COMMENT Original source text: Human DNA.  
A cytosine to thymidine transition at nucleotide position 868 results in substitution of TRP-240 for ARG-240.  
Location/Qualifiers  
1..1742  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/haplotype="OKT4-negative"  
76..1452  
/gene="CD4"  
76..1452  
/gene="CD4"  
/note="T4 surface glycoprotein precursor"  
/codon\_start=1  
/protein\_id="AAA16069.1"  
/db\_xref="GI:179144"  
/translation="MNRGVPRHLLVLQALLPAATQGRKVLGKKDVTVEITCTAS  
QKSIQPHMKNNOIKILGNQSFITGSPKLNDRADSRSLMDQNPPLIKNLKIE  
DSPTVCEVEDQKEEVLVFLGTANSDFLILQOSLTLSPSSPSVOCRRSPRG  
KNIOGGKTLVSQLELDGSMCTVIONOKVEFKIDIVLAFORASSIVYKKEGEO  
VERSPPLATFVERKLTGSGELMOWAERASSKSMITPDLKNKESYKRVTDPKIQMK  
KLPHILTPQALPQYAGSNLTLALEKTKLHQEVNLVVRATQLOKNTLCEWGT  
SPKLMISLKLKNEKAVSKREKAVWVNLNPAQMWOCILSDSQVLLSNIKVLPTWST  
PVQPMALIVLGVALILFLIGLIFFCVCRHRRROAERMSQIKRLSEKKTCCPBR  
FOKTCSPPI"  
76..150  
/gene="CD4"  
/note="T4 surface glycoprotein signal peptide"  
151..1449  
/gene="CD4"  
/product="T4 surface glycoprotein mature peptide"  
792  
/gene="CD4"  
/note="g in OKT4 epitope deficiency; c in wt"

ORIGIN  
Query Match 43.4%; Score 624.4; DB 9; Length 1742;  
Best Local Similarity 99.8%; Pred. No. 1.8e-134;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 ATGAAACCGGGAGAGTCCCTTTAGGCACTTGTGGTGTGCAACTGGCGCTCTCCCA 60  
76 ATGAACCGGGAGAGTCCCTTTAGGCACTTGTGGTGTGCAACTGGCGCTCTCCCA 135  
61 GCAGCCACTCAGGAAAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAACCTGACC 120  
136 GCAGCCACTCAGGAAAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAACCTGACC 195  
121 TGTACAGCTTCCAGAAAGAGATCAATTCACAGTGAACCACTCCAGATTAAG 180

Db 196 TGTACAGCTTCCCAAGAAAGACATTCATCCACTGGAAGAACTCCAAACGATTAAG 255  
QY 181 ATCTGGGAATCAGGGCTCCTTTACTAAAGATCCAAAGCTGAATGATCGGCT 240  
Db 256 ATCTGGGAATCAGGGCTCCTTTACTAAAGATCCAAAGCTGAATGATCGGCT 315  
QY 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 300  
Db 316 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 375  
QY 301 ATGAAAGCTCAGATCTTACATCTGTGAAGTGAAGACCAAGAAAGGATGCAATTG 360  
Db 376 ATGAAAGCTCAGATCTTACATCTGTGAAGTGAAGACCAAGAAAGGATGCAATTG 435  
QY 361 CTAGTGTGGATTTAGCTGCAACTCTGACACCACTGCTTCAAGGGGACAGCTGACC 420  
Db 436 CTAGTGTGGATTTAGCTGCAACTCTGACACCACTGCTTCAAGGGGACAGCTGACC 495  
QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGCT 480  
Db 496 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGCT 555  
QY 481 AAAAACAATCAGGGGGGAAAGACCTTCTGCTGTCTGAGCTGAGATGATGAGC 540  
Db 556 AAAAACAATCAGGGGGGAAAGACCTTCTGCTGTCTGAGCTGAGATGATGAGC 615  
QY 541 ACCTGACATGCACTGTCTTGAAGACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 600  
Db 616 ACCTGACATGCACTGTCTTGAAGACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 675  
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
Db 676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701

RESULT 8  
LOCUS 108116 1910 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 3 from Patent EP 0325262.  
ACCESSION 108116  
VERSION 108116.1 GI:589172  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1910)  
AUTHORS Brian,S.D.  
TITL Cloned genes encoding IG-CD4 fusion proteins and the use thereof  
JOURNAL Patent: EP 0325262-A2 3 26-JUL-1989;  
FEATURES  
source 1..1910  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 43.4%; Score 624.4; DB 6; Length 1910;  
Best Local Similarity 99.8%; Pred. No. 1.8e-134;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGCTGCACTGGCGCTCTCCCA 60  
Db 111 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGCTGCACTGGCGCTCTCCCA 170  
QY 61 GCAGCCACTCAGGAAAGAAAGTGTGCTGGCAAAAAGGGGATACAGTGAACCTGACC 120  
Db 171 GCAGCCACTCAGGAAAGAAAGTGTGCTGGCAAAAAGGGGATACAGTGAACCTGACC 230  
QY 121 TGTACAGCTTCCCAAGAAAGAGCATATTCATCTGGAATAAATCTCAACAGATTAAG 180  
Db 231 TGTACAGCTTCCCAAGAAAGAGCATATTCATCTGGAATAAATCTCAACAGATTAAG 290  
QY 181 ATCTGGGAATCAGGGCTCCTTTACTAAAGATTCATCAAGATGATCGGCT 240

Db 291 ATCTGGGAATCAGGGCTCCTTTACTAAAGTCCATCCAACTGATATATGCGCT 350  
QY 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTTAAG 300  
Db 351 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTTAAG 410  
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAAAGGATGCAATTG 360  
Db 411 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAAAGGATGCAATTG 470  
QY 361 CTAGTGTGGATTTAGCTGCAACTCTGACACCACTGCTTCAAGGGGACAGCTGACC 420  
Db 471 CTAGTGTGGATTTAGCTGCAACTCTGACACCACTGCTTCAAGGGGACAGCTGACC 530  
QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGCT 480  
Db 531 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGCT 590  
QY 481 AAAAACAATCAGGGGGGAAAGACCTTCTGCTGTCTGAGCTGGAAGTCCAGATAGTGC 540  
Db 591 AAAAACAATCAGGGGGGAAAGACCTTCTGCTGTCTGAGCTGGAAGTCCAGATAGTGC 650  
QY 541 ACTTGACATGCACTGTCTTGAAGACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 600  
Db 651 ACTTGACATGCACTGTCTTGAAGACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 710  
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
Db 711 GTGCTAGCTTTCCAGAAAGGCTCCAG 736

RESULT 9  
LOCUS 108115 3133 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent EP 0325262.  
ACCESSION 108115  
VERSION 108115.1 GI:589171  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3133)  
AUTHORS Brian,S.D.  
TITL Cloned genes encoding IG-CD4 fusion proteins and the use thereof  
JOURNAL Patent: EP 0325262-A2 1 26-JUL-1989;  
FEATURES  
source 1..3133  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 43.4%; Score 624.4; DB 6; Length 3133;  
Best Local Similarity 99.8%; Pred. No. 1.8e-134;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGCTGCACTGGCGCTCTCCCA 60  
Db 111 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGCTGCACTGGCGCTCTCCCA 170  
QY 61 GCAGCCACTCAGGAAAGAAAGTGTGCTGGCAAAAAGGGGATACAGTGAACCTGACC 120  
Db 171 GCAGCCACTCAGGAAAGAAAGTGTGCTGGCAAAAAGGGGATACAGTGAACCTGACC 230  
QY 121 TGTACAGCTTCCCAAGAAAGAGCATATTCATCTGGAATAAATCTCAACAGATTAAG 180  
Db 231 TGTACAGCTTCCCAAGAAAGAGCATATTCATCTGGAATAAATCTCAACAGATTAAG 290  
QY 181 ATCTGGGAATCAGGGCTCCTTTACTAAAGTTCATCAAGATGATCGGCT 240  
Db 291 ATCTGGGAATCAGGGCTCCTTTACTAAAGTTCATCAAGATGATCGGCT 350  
QY 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTTAAG 300

Db	351	GACTCAAGAAAGGCTTTGGGACCAAGGAACTTCCCCCTGATCATCAAGATCTTTAAG	410
Oy	301	ATAGAAAGCTCAGATATCTTAACTCTGTGAAGTGAGAGACCAAGAGAGAGGTGCATTTG	360
Db	411	ATAGAAAGCTCAGATATCTTAACTCTGTGAAGTGAGAGACCAAGAGAGAGGTGCATTTG	470
Oy	361	CTAGGTTCGGATTTGACCTGCCAACCTGTACACCCCACTGCTTCAGGGGCACAGCCTGACC	420
Db	471	CTAGGTTCGGATTTGACCTGCCAACCTGTACACCCCACTGCTTCAGGGGCACAGCCTGACC	530
Oy	421	CTGACCTTGGAGAGGCCCCCTGTGTGTGTAGCCCCCTCAGTGCATGTAGAGATCCAAAGGGGT	480
Db	531	CTGACCTTGGAGAGGCCCCCTGTGTGTGTGTAGCCCCCTCAGTGCATGTAGAGATCCAAAGGGGT	590
Oy	481	AAAAACATACAGGGGGGGGAGAGACCCCTCTCCGTGTCTCAGCTGAGAGCTCAGAGATAGTGGC	540
Db	591	AAAAACATACAGGGGGGGGAGAGACCCCTCTCCGTGTCTCAGCTGAGAGCTCAGAGATAGTGGC	650
Oy	541	ACCTGAGATGCACTGTCTTTCGACAGAACCAAGAAAGGTGTGAGTTAAATATGACATCGTG	600
Db	651	ACCTGAGATGCACTGTCTTTCGACAGAACCAAGAAAGGTGTGAGTTCAATATGACATCGTG	710
Oy	601	GTGCTAGCTTCCAGAAAGGCTCCGGG	626
Db	711	GTGCTAGCTTCCAGAAAGGCTCCGGG	736
RESULT 10			
LOCUS	AY438650/c	7533 bp	DNA linear SYN 11-NOV-2003
DEFINITION	Cloning vector pMACSiBac, complete sequence.		
ACCESSION	AY438650		
VERSION	AY438650.1	GI:38197734	
KEYWORDS	Cloning vector pMACSiBac		
SOURCE	Cloning vector pMACSiBac		
ORGANISM	other sequences; artificial sequences; vectors.		
REFERENCE	1 (bases 1 to 7533)		
AUTHORS	Phillips,B., Foretner,M. and Mayr,L.M.		
TITLE	A baculovirus expression system for magnetic sorting of infected cells and enhanced titer determination		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 7533)		
AUTHORS	Phillips,B., Foretner,M. and Mayr,L.M.		
TITLE	Direct Submision		
JOURNAL	Submitted (16-OCT-2003) LDC / BSOI, Novartis Pharma AG, MSU-88.601, Basel CH-4002, Switzerland		
FEATURES	Location/Qualifiers		
source	1..7533		
	/organism="Cloning vector pMACSiBac"		
	/mol_type="other DNA"		
	/db_xref="taxon:254251"		
gene	689..1549		
	/gene="bla"		
CDs	689..1549		
	/gene="bla"		
	/codon_start=1		
	/transl_table=11		
	/product="beta-lactamase"		
	/protein_id="AA13800.1"		
	/db_xref="GI:38197735"		
	/translation="MSIOHPRVALIPPEAFCLPVFAHPETLVKVKADQLGARVGGV		
	IELDNSKILTESFPERPMPMSPTFKYLCAVLSRIDAOEOLGRIRHYSNDLVLE		
	YSPVTEKILDTGMVTRKELCSAATMSDNTANILITTGKPELTAFLAHMNGDHVTE		
	DRWEPELNEALPNDERDPTTMYAAMTTRIKLLTGLSLTLASRQQLIDMEADKTAGP		
	LRSLAPAGMFLADRSAGAEKRSRGIILALGPDGKPSRIVLVITTSQATIMDERNRQILR		
	EIGSLALPQHW"		
	complement (2302..3435)		
gene	/gene="aacC1"		
	complement (2302..3435)		
CDs	/gene="aacC1"		
	/codon_start=1		

```

gene
CDS
        /transl_table=11
        /product="gentamicin-3'-acetyl transferase"
        /protein_id="AAR3801.1"
        /db_xref="GI:38197736"
        /translation="MLRSSNDVYQQGSRPRTKLGGSGWGIIRTCRLGDPGVKSNRAAL
        DLGREGEDVATYSOHQPDSDYLGNLIRSTFTALAFDGEAVGALAAVLPREFEOP
        RSEIYIYDLAVSGEHRROGIATLAILNLKHEANLAGYIVQADVDGDDPAVALYTKL
        GIREVWHFDIPSTAT"
        complement(4301..6394)
        /gene="CD4-EGFP"
        complement(4301..6394)
        /gene="CD4-EGFP"
        complement(4301..6394)
        /note="CD4 and enhanced GFP fusion protein"
        /codon_start=1
        /transl_table=11
        /product="CD4-EGFP"
        /protein_id="AAR3802.1"
        /db_xref="GI:38197737"
        /translation="MNRGVPFRHLLVLTQLALPAATQGRKVVLAGKKGDTVELTCTAS
        OKSIOFHMNSNOIKILGNGSELTGPKLNDRAISRSLDNGGFPILIKNLKE
        DSPTYCEVDOKKEVOLIVGLTANDDTLLOGSLDITLSESPSPVOCSPRG
        KTIQGGKTLISVQLDELDSCTITCTYIQNOKRPEKDIIVLAFQKASLIVYKKEGQ
        VESFPLATVEKLTSGELMWAERSSKSLITPLKNKSVKRVATDPPKLQKPK
        KLPLHLTLPALPOYAGSNLTALAEKAEKSLKQELVNRATQLOKNTLCEWGPST
        SPTMLSTLKEKAEKSKREKAVVWVLPNPAWMOCLSDSGVLLESNTKVLPTWST
        PVPOMALIVLGVAGLLPIGLGIFPCVRCHRRROAERMSQIKRLSEKTKCQCPHR
        FORTKPSIVSKGBELFTGVVPIIVLVEDGPNVGHKFSVSGEGDATTYGLTKLFTCT
        TGKLPVWPLTLLTITLYGQCPSRYPDHMKOHPFKSAMEGVQERTIPKIDGNVKK
        TRABVKEGDTLVNRIELKQIDFKEDGNITLGHLEIYNASHNYITADKQNGIKNVF
        KIRHNIIDGVSVALDHYQONTPIGDGVLLIPDNHYLSTQSALSDPKEKRDHVLLEF
        VTAAIGHIDGDELADLYK"

```

## ORIGIN

Query Match	Score	DB	Length
43.4%	624.4	12	7533

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTGCTTCTGGTCTGCAACTGGCGCTCCTCCA 60

Db 6394 ATGAACCGGGAGTCCCTTTAGGCACTTGCTTCTGGTGTGCAACTGGCGCTCCTCCA 6335

61 GCAGCCACTCAGGGAAGAAGTGGTCTGGCAGGATACAGTGAAGTACC 120

Db 6334 GCAGCCACTCAGGGAAGAAGTGGTCTGGGCAAAAAGGGATACAGTGGAAGTACC 6275

121 TGTACAGCTTCCAGAGAGAGCATACATTTCCA CTGGAAAACTCCAAC CAGATAAG 180

Db 6274 TGTACAGCTTCCAGAGAAGACATACAATTCCACTGGAAAACTCCAACGATTAAG 6215

181 ATTCGGAATCAGGGCTCCTTCTTAACCTAAAGGTCATCCAGCTGAATGATCGGCT 240

Db 6214 ATTCTGGGAATCAGGCTCCTTCTTAAGGTCATCCAAGCTGAATGATCGCGCT 6155

241 GACTCAGAAGAGCCTTTGGGACCAAGAACTTCCCCCTGATCATCAAGAATCTTAAG 300

Db 6154 GACTCAGAGAAGCCTTGGGACCAAGAACTTCCCCCTGATCATCAAGAATCTTAAG 60955

301 ATAGAGACTCAGATACTTACATCTGTGAGTGGAGGACCAGAAGGAGGAGGTCAATTG 360

Db 6094 ATAGAGACTCAGATACTTACATCTGTGAGTGGAGGACCAGAGGAGGAGGTGCAATTG 6035

QY 361 CTAGTGTTCGATTGACTGCCACTCTGACACCCACCTGCTTCAGGGGCAGAGCCTGACC 420

Db 6034 CTAGTGTTCGATTGACTGCCAACTCTGACACCCCACTGCTTCAGGGGCAGAGCCTGACC 5975

421 CTGACCTTGAGAGCCCCCTGTAAGGCCCTCAGTGCAATGTAGAGTCCAAGGGT 480

Db 5974 CTGACCTTGAGAGCCCCCTGGTAGGCCCCCTCAGTGCATGTAGAGTCCAAGGGT 5915

481 AAAAATACAGGGGGAGACCTCTCCGTCTCAGCTGAGCTCCAGATAGTGGC 540

Db 5914 AAAAATATACAGGGGGAGACCCTCTCCGTGTCTCAGCTGAGCTCCAGGATAGTGGC 5855

OY	541	ACCTGGACATCATCTGTTTCAGAAACGAAAGGTGGAGTTCAAAATGACATCGTG	600
Db	5554	ACCTGGACATCATCTGTTTCAGAAACGAAAGGTGGAGTTCAAAATGACATCGTG	5795
OY	601	GTGCTAGCTTTCAGAAAGGCTCCGG	626
Db	5794	GTGCTAGCTTTCAGAAAGGCTCCAG	5769
RESULT	11		
LOCUS	106223	1273 bp	DNA
DEFINITION	Sequence 1 from Patent EP 0313377.	linear	PAT 02-DEC-1994
VERSION	106223		
KEYWORDS	106223.1	GI:590426	
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1273)		
TITLE	Deen, R.C., Foleina-Wasserman, G.M., Inacker, R.H. and Sweet, R.W.		
JOURNAL	Process for purification of soluble T4		
FEATURES	Patent: EP 0313377-A1 1 26-APR-1989;		
source	location/qualifiers		
	1..1273		

ORIGIN	/mol_type="unassigned DNA"				
Query Match	43.2%	Score 622.8;	DB 6;	Length 1273;	
Best Local Similarity	99.7%;	Pred. No. 4.2e-134;			
Matches 624;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;	
QY	1	ATGAACCGGGGAGTCCCTTTAAGCACTTGCTCGGTGCTGCAACTGGGCTCTCCCA	60		
Db	76	ATGAACCGGGGAGTCCCTTTAAGCACTTGCTCGGTGCTGCAACTGGGCTCTCCCA	135		
QY	61	GCAGCCACTCAGGAAAGAAAGTGCTGTGGGCAAAAAGGGATATCAGTGAATGACC	120		
Db	136	GCAGCCACTCAGGAAAGAAAGTGCTGTGGGCAAAAAGGGATATCAGTGAATGACC	195		
QY	121	TGTACAGCTTCCCAAGAGAAGCATATCAATTCCACTGGAAAACTCCACAGATAAG	180		
Db	196	TGTACAGCTTCCCAAGAGAAGCATATCAATTCCACTGGAAAACTCCACAGATAAG	255		
QY	181	ATTCTGGGAATCAGGGCTCCTTCTTAATAAGTCCATCAAGCTGAATGATGGGCT	240		
Db	256	ATTCTGGGAATCAGGGCTCCTTCTTAATAAGTCCATCAAGCTGAATGATGGGCT	315		
QY	241	GACTCAAGAAAGACCTTTTGGGACCAAGGAATCTTCCCTGATCATCAAGATCTTAA	300		
Db	316	GACTCAAGAAAGACCTTTTGGGACCAAGGAATCTTCCCTGATCATCAAGATCTTAA	375		
QY	301	ATAGAAGCTCAGATACTTACATCTGTGAAGTGAAGACAGAAAGAGAGTGCATTTG	360		
Db	376	ATAGAAGCTCAGATACTTACATCTGTGAAGTGAAGACAGAAAGAGAGTGCATTTG	435		
QY	361	CTAGTGTTCGGAATTGACGTGCCAATCTTGACACCCACTGCTTCAAGGGGCAAGCTGACC	420		
Db	436	CTAGTGTTCGGAATTGACGTGCCAATCTTGACACCCACTGCTTCAAGGGGCAAGCTGACC	495		
QY	421	CTGACCTTTGGAGAGCCCCCTGTGTATGTAGCCCTCAGTGCATATAGGAGTCCAAAGGGAT	480		
Db	496	CTGACCTTTGGAGAGCCCCCTGTGTATGTAGCCCTCAGTGCATATAGGAGTCCAAAGGGAT	555		
QY	481	AAAAACATACAGGGGGGGAGAACCTCTTCGCTGTCTCACTGAGACTCAGGAATAGTGC	540		
Db	556	AAAAACATACAGGGGGGGAGAACCTCTTCGCTGTCTCACTGAGACTCAGGAATAGTGC	615		
QY	541	ACCTGGAATGCACTGCTTTGCAAGAACCAAGAAAGGTGAGATTCAAAATATGACATGCTG	600		
Db	616	ACCTGGAATGCACTGCTTTGCAAGAACCAAGAAAGGTGAGATTCAAAATATGACATGCTG	675		

Oy	601	GTGCTAGCTTCCAGAAGCCTCCGG	626
Dd	676	GTGCTAGCTTCCAGAAGCCTCCAG	701
RESULT 12			
LOCUS	107147	107147	1273 bp DNA linear PAT 02-DEC-1994
DEFINITION	Sequence 1 from Patent EP 0331356.		
ACCESSION	107147		
VERSION	107147.1	GI:590041	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1273)		
TITLE	Arthos,J., Clark,P.E., Fornwald,J.A., Brawner,M.E., Deen,K.C.,		
JOURNAL	Gorman,J.A., Sachse,G.M., Sweet,R.W. and Taylor,D.P.		
FEATURES	Expression of HIV binding proteins		
	Patent: EP 0331356-A2 1 06-SEP-1989;		
	Location/Qualifiers		
source	1..1273		
	/organism="unknown"		
ORIGIN	/mol_type="unassigned DNA"		

Query Match	43.2%	Score 622, 8; DB 6;	Length 1273;
Best Local Similarity	99.7%	Pred. No. 4,2e-134;	
Matches 624;	Conservative 0;	Mismatches 2;	Indels 0; Gaps 0;

QY	1	ATGACCGGGGAGTCCCTTTTAGCCACTGCTTGTGTGCTGCACTGCGCTCTCTCCCA	60
Db	76	ATGAAACCGGGGAGTCCCTTTTAGCCACTGTGCTTGTGTGCTGCACTGCGCTCTCTCCCA	135
QY	61	GCACCCACTCAGGGGAAAGAAAGTGTGTCTGGGCAAAAAGGGGATACAGTGGAACTGACC	120
Db	136	GCACCCACTCAGGGGAAAGAAAGTGTGTCTGGGCAAAAAGGGGATACAGTGGAACTGACC	195
QY	121	TGTACAGCTTCCCAAGAAAGAGCATCAATTCCACTGGAAAACTCCAAACAGATAAAG	180
Db	196	TGTACAGCTTCCCAAGAAAGAGCATCAATTCCACTGGAAAACTCCAAACAGATAAAG	255
QY	181	ATTGTGGGAAATCGGGGCTCCTTTTAACTTAAAGTGCATCCAACTGAATGATCGCGCT	240
Db	256	ATTGTGGGAAATCGGGGCTCCTTTTAACTTAAAGTGCATCCAACTGAATGATCGCGCT	315
QY	241	GACTCAAGAAAGAACCTTTTGGGACCAAGAAACTTCCCTCGATCATCAAGATCTTAAAG	300
Db	316	GACTCAAGAAAGAACCTTTTGGGACCAAGAAACTTCCCTCGATCATCAAGATCTTAAAG	375
QY	301	ATAGAAGACTCAGATCTTACATCTGTGAAGTGAGAGACCAAGAGAGAGGTGCATTG	360
Db	376	ATAGAAGACTCAGATCTTACATCTGTGAAGTGAGAGACCAAGAGAGAGGTGCATTG	435
QY	361	CTAGTGTGGATTGACTGCGCAACTCTGACACCCACCTGCTTCAAGGGCAAGGCTGACC	420
Db	436	CTAGTGTGGATTGACTGCGCAACTCTGACACCCACCTGCTTCAAGGGCAAGGCTGACC	495
QY	421	CTGACCTTGGAGAGCCCCCTGTGTAGTAGCCCTCAGTGCATATGAGAGTCCAAAGGGT	480
Db	496	CTGACCTTGGAGAGCCCCCTGTGTAGTAGCCCTCAGTGCATATGAGAGTCCAAAGGGT	555
QY	481	AAAAACATACAGGGGGGGGAAAGACCTCTCCGTGTCTCAAGTGGAGCTTCAGATATGTGGC	540
Db	556	AAAAACATACAGGGGGGGGAAAGACCTCTCCGTGTCTCAAGTGGAGCTTCAGATATGTGGC	615
QY	541	ACCTGGAATGACATGTCTTTCAGAAACCAAGAAAGAGTGGAGTTCAAAATATGACATCGTG	600
Db	616	ACCTGGAATGACATGTCTTTCAGAAACCAAGAAAGAGTGGAGTTCAAAATATGACATCGTG	675
QY	601	GTGCTAGCTTTCAGAAAGGCTCCGG	626
Db	676	GTGCTAGCTTTCAGAAAGGCTCCGAG	701



RESULT 13  
 LOCUS AR067943 1304 bp DNA linear PAT 29-SEP-1999  
 DEFINITION Sequence 28 from patent US 5851828.  
 ACCESSION AR067943  
 VERSION AR067943.1 GI:5999165  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE Unclassified.  
 AUTHORS 1 (bases 1 to 1304)  
 TITLE Seed, B., Banapour, B., Romeo, C. and Kolanus, W.  
 JOURNAL Targeted cytolysis of HIV-infected cells by chimeric CD4  
 FEATURES receptor-bearing cells  
 Patent: US 5851828-A 28 22-DEC-1998;  
 Location/Qualifiers  
 1..1304  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN  
 Query Match 43.2%; Score 622.8; DB 6; Length 1304;  
 Best Local Similarity 99.7%; Pred. No. 4.2e-134;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAACCGGGAGTCCCTTTTAAAGCACTTGTGCTGCGCACTGGCGCTCTCCCA 60  
 111 ATGAACCGGGAGTCCCTTTTAAAGCACTTGTGCTGCGCACTGGCGCTCTCCCA 170  
 61 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAAGGGATACAGTGAATGACC 120  
 171 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAAGGGATACAGTGAATGACC 230  
 121 TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGGAATACTCCACAGATTAAG 180  
 231 TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGGAATACTCCACAGATTAAG 290  
 181 ATTCTGGGAATCAGGGCTCTCTTTAACTTAAGTCCATTCAGTGAATGATCGGCT 240  
 291 ATTCTGGGAATCAGGGCTCTCTTTAACTTAAGTCCATTCAGTGAATGATCGGCT 350  
 241 GACTCAAGAAAGAGCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 300  
 351 GACTCAAGAAAGAGCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 410  
 301 ATAGAAGACTCAGATCTTAATCATCTGTGAAGTGAAGACCAAGAGAGGTGCAATTG 360  
 411 ATAGAAGACTCAGATCTTAATCATCTGTGAAGTGAAGACCAAGAGAGGTGCAATTG 470  
 361 CTAGTGTGGATTTGACTGCGCAACTGTGACACCACTGCTTCAGGGGCAAGCTTACC 420  
 471 CTAGTGTGGATTTGACTGCGCAACTGTGACACCACTGCTTCAGGGGCAAGCTTACC 530  
 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGAATGTAGAGATCCAAAGGGT 480  
 531 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGAATGTAGAGATCCAAAGGGT 590  
 481 AAAAACAATACAGGGGGGAAAGCCCTCTCGTGTCTCAGTGAAGCTCCAGATAGTGGC 540  
 591 AAAAACAATACAGGGGGGAAAGCCCTCTCGTGTCTCAGTGAAGCTCCAGATAGTGGC 650  
 541 ACCTGACATGCACTGTCTTTCAGAAACCAAGAAAGGTGAGATTCAAAATAGACTGTG 600  
 651 ACCTGACATGCACTGTCTTTCAGAAACCAAGAAAGGTGAGATTCAAAATAGACTGTG 710  
 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
 711 GTGCTAGCTTTCCAGAAAGGCTCCAG 736

RESULT 14

ARI66802  
 LOCUS ARI66802 1304 bp DNA linear PAT 17-OCT-2001  
 DEFINITION Sequence 28 from patent US 6284240.  
 ACCESSION ARI66802  
 VERSION ARI66802.1 GI:16243143  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE Unclassified.  
 AUTHORS 1 (bases 1 to 1304)  
 TITLE Seed, B., Banapour, B., Romeo, C. and Kolanus, W.  
 JOURNAL Targeted cytolysis of HIV-infected cells by chimeric CD4  
 FEATURES receptor-bearing cells  
 Patent: US 6284240-A 28 04-SEP-2001;  
 Location/Qualifiers  
 1..1304  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN  
 Query Match 43.2%; Score 622.8; DB 6; Length 1304;  
 Best Local Similarity 99.7%; Pred. No. 4.2e-134;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAACCGGGAGTCCCTTTTAAAGCACTTGTGCTGCGCACTGGCGCTCTCCCA 60  
 111 ATGAACCGGGAGTCCCTTTTAAAGCACTTGTGCTGCGCACTGGCGCTCTCCCA 170  
 61 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAAGGGATACAGTGAATGACC 120  
 171 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAAGGGATACAGTGAATGACC 230  
 121 TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGGAATACTCCACAGATTAAG 180  
 231 TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGGAATACTCCACAGATTAAG 290  
 181 ATTCTGGGAATCAGGGCTCTCTTTAACTTAAGTCCATTCAGTGAATGATCGGCT 240  
 291 ATTCTGGGAATCAGGGCTCTCTTTAACTTAAGTCCATTCAGTGAATGATCGGCT 350  
 241 GACTCAAGAAAGAGCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 300  
 351 GACTCAAGAAAGAGCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 410  
 301 ATAGAAGACTCAGATCTTAATCATCTGTGAAGTGAAGACCAAGAGAGGTGCAATTG 360  
 411 ATAGAAGACTCAGATCTTAATCATCTGTGAAGTGAAGACCAAGAGAGGTGCAATTG 470  
 361 CTAGTGTGGATTTGACTGCGCAACTGTGACACCACTGCTTCAGGGGCAAGCTTACC 420  
 471 CTAGTGTGGATTTGACTGCGCAACTGTGACACCACTGCTTCAGGGGCAAGCTTACC 530  
 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGAATGTAGAGATCCAAAGGGT 480  
 531 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGAATGTAGAGATCCAAAGGGT 590  
 481 AAAAACAATACAGGGGGGAAAGCCCTCTCGTGTCTCAGTGAAGCTCCAGATAGTGGC 540  
 591 AAAAACAATACAGGGGGGAAAGCCCTCTCGTGTCTCAGTGAAGCTCCAGATAGTGGC 650  
 541 ACCTGACATGCACTGTCTTTCAGAAACCAAGAAAGGTGAGATTCAAAATAGACTGTG 600  
 651 ACCTGACATGCACTGTCTTTCAGAAACCAAGAAAGGTGAGATTCAAAATAGACTGTG 710  
 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
 711 GTGCTAGCTTTCCAGAAAGGCTCCAG 736

RESULT 15  
 AR560088 1304 bp mRNA linear PAT 08-OCT-2004  
 LOCUS AR560088  
 DEFINITION Sequence 28 from patent US 6753162.

ACCESSION ARS60088  
VERSION ARS60088.1 GI:53970428  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 1304)  
AUTHORS Seed, B., Banapour, B., Romeo, C. and Kolanus, W.  
TITLE Targeted cytolysis of HIV-infected cells by chimeric CD4  
receptor-bearing cells  
JOURNAL Patent: US 6753162-A 28 22-JUN-2004;  
FEATURES location/Qualifiers  
source 1..1304  
/organism="unknown"  
/mol\_type="mRNA"

## ORIGIN

Query Match 43.2%; Score 622.8; DB 6; Length 1304;  
Best Local Similarity 99.7%; Pred.No. 4.2e-134;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGACCGGGAGTCCCTTTAGCACTGCTGCTGCTGCAACTGGCGCTCCTCCA 60  
DB |||  
QY 111 ATGACCGGGAGTCCCTTTAGCACTGCTGCTGCTGCAACTGGCGCTCCTCCA 170  
DB |||  
QY 61 GCAGCCACTCAGGGAAACAAAGTGCTGCGCAAAAAGGGGATACAGTGAACTGACC 120  
DB |||  
QY 171 GCAGCCACTCAGGGAAACAAAGTGCTGCGCAAAAAGGGGATACAGTGAACTGACC 230  
DB |||  
QY 121 TGTACAGCTCCCGAAGAAAGACATCAATTCACCTGGAATACTCCACCAATAAG 180  
DB |||  
QY 231 TGTACAGCTCCCGAAGAAAGACATCAATTCACCTGGAATACTCCACCAATAAG 290  
DB |||  
QY 181 ATTCTGGGAATCAGGCTCCTTCTTAAGTCAATCAAGTCAAGTCAATGATCGGCT 240  
DB |||  
QY 291 ATTCTGGGAATCAGGCTCCTTCTTAAGTCAATCAAGTCAAGTCAATGATCGGCT 350  
DB |||  
QY 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 300  
DB |||  
QY 351 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 410  
DB |||  
QY 301 ATGAAGACTCAGTACTTCACTGTAAGTGGAGACCAAGAGAGAGGTCATTTG 360  
DB |||  
QY 411 ATGAAGACTCAGTACTTCACTGTAAGTGGAGACCAAGAGAGAGGTCATTTG 470  
DB |||  
QY 361 CTAGTGTTCGGATTGACTGCACTTGACACCCAGCTGCTTCAGGGCAGAGCTGACC 420  
DB |||  
QY 471 CTAGTGTTCGGATTGACTGCACTTGACACCCAGCTGCTTCAGGGCAGAGCTGACC 530  
DB |||  
QY 421 CTGACCTTGAAGAGCCCTCTGTAGTACCCCTCAGTGCATGTAGAGTCCAAAGGGT 480  
DB |||  
QY 531 CTGACCTTGAAGAGCCCTCTGTAGTACCCCTCAGTGCATGTAGAGTCCAAAGGGT 590  
DB |||  
QY 481 AAAAACAATACAGGGGGGGAAGACCCCTCCGTCTCAGCTGGAGCTCCAGGATAGTGGC 540  
DB |||  
QY 591 AAAAACAATACAGGGGGGGAAGACCCCTCCGTCTCAGCTGGAGCTCCAGGATAGTGGC 650  
DB |||  
QY 541 ACCTGACATGACTGTCTTGCAAGAACAGAAAGGTGAGTTCAAAATAGACATGCTG 600  
DB |||  
QY 651 ACCTGACATGACTGTCTTGCAAGAACAGAAAGGTGAGTTCAAAATAGACATGCTG 710  
DB |||  
QY 601 GTGCTAGCTTTCCAGAAAGCCTCCGG 626  
DB |||  
QY 711 GTGCTAGCTTTCCAGAAAGCCTCCAG 736  
DB |||

Search completed: September 20, 2005, 10:17:33  
Job time : 6463 secs



PT HIV neutralization, comprises two different binding domains, inducing-  
 PT binding domain and induced-binding domain functionally linked by linker.  
 PS Claim 36; Page 47-48; 55pp; English.  
 XX

CC sCD4-sCFv(17b) is a neutralizing bispecific fusion protein capable of  
 CC binding to two sites of its target protein. The protein comprises a first  
 CC binding domain capable of binding to an inducing site on the target  
 CC protein, a second binding domain capable of forming neutralising complex  
 CC with an induced epitope of the target protein and a linker connecting the  
 CC (containing domains D1 and D2) fused to a single chain Fv portion of  
 CC antibody 17b via a linker. sCD4-sCFv(17b), its variant, analogue or  
 CC derivative is used for inactivating gp120 protein of HIV, and for  
 CC neutralising HIV. It is also used for blocking and preventing the binding  
 CC of the viral or recombinant gp120 protein to soluble CD4 or lymphocyte  
 CC CD4 and for inhibiting HIV replication. The chimeric proteins is  
 CC therefore useful for treating HIV infection and also AIDS. It is are  
 CC particularly useful in the prevention of infection during or immediately  
 CC after HIV exposure (e.g., mother/infant transmission, post-exposure  
 CC prophylaxis, and as a topical inhibitor) and for providing long term  
 CC resistance to HIV infections and AIDS. Gene therapy is used to secrete  
 CC the bispecific protein at mucosal surfaces, such as the vaginal, rectal  
 CC or oral mucosa. The fusion proteins is highly potent, broadly cross-  
 CC reactive with neutralising antibody with high in vivo activity and no Fc-  
 CC mediated undesirable targeting properties. When the fusion protein is  
 CC substantially derived from human proteins, it has minimal immunogenicity  
 CC and toxicity in humans which is of great value in prevention of infection  
 CC during or immediately after HIV exposure  
 XX

Sequence 1440 BP; 345 A; 354 C; 452 G; 289 T; 0 U; 0 Other;

Query Match 100.0%; Score 1440; DB 3; Length 1440;

Best Local Similarity 100.0%; Pred. No. 1.5e-311;

Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAACCGGGAGTCCCTTTTAAAGCACTTGTGCTGCAACTGGCGCTCTCCCA 60  
 DB 1 ATGAAACCGGGAGTCCCTTTTAAAGCACTTGTGCTGCAACTGGCGCTCTCCCA 60  
 QY 61 GCAGCCACTCAGGAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC 120  
 DB 61 GCAGCCACTCAGGAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC 120  
 QY 121 TGTACAGCTTCCAGAAAGAGCATTAATTCACCTGGAATACTCAACAGATTAAG 180  
 DB 121 TGTACAGCTTCCAGAAAGAGCATTAATTCACCTGGAATACTCAACAGATTAAG 180  
 QY 181 ATTCTGGGAATCAGGGCTCTTTTAACTTAAGTCAATCCAGTGAATGATCGGCT 240  
 DB 181 ATTCTGGGAATCAGGGCTCTTTTAACTTAAGTCAATCCAGTGAATGATCGGCT 240  
 QY 241 GACTCAAGAAAGCTTTGGGACCAAGAACTTCCCGTGAATCATCAAGAAATCTTAAG 300  
 DB 241 GACTCAAGAAAGCTTTGGGACCAAGAACTTCCCGTGAATCATCAAGAAATCTTAAG 300  
 QY 301 ATAGAAAGCTCAGTACTTATCATCTGTGAAGTGAAGCAAGAAAGAGGTGCAATTG 360  
 DB 301 ATAGAAAGCTCAGTACTTATCATCTGTGAAGTGAAGCAAGAAAGAGGTGCAATTG 360  
 QY 361 CTAGTGTTCGGATTGATGCAACTTGACCACTGCTTCAAGGGGCAAGGCTTAAC 420  
 DB 361 CTAGTGTTCGGATTGATGCAACTTGACCACTGCTTCAAGGGGCAAGGCTTAAC 420  
 QY 421 CTGACCTTTGGAGAGCCCCCTGATAGAGCCCTCACTGATGCAATATAGAGGAGT 480  
 DB 421 CTGACCTTTGGAGAGCCCCCTGATAGAGCCCTCACTGATGCAATATAGAGGAGT 480  
 QY 481 AAAAACAATACAGGGGGGAAAGCCCTCTCGTGTCTCAGCTGAGCTCCAGATAGTGC 540  
 DB 481 AAAAACAATACAGGGGGGAAAGCCCTCTCGTGTCTCAGCTGAGCTCCAGATAGTGC 540  
 QY 541 ACCTGGACATGCACTGTCTTGCAAGAACCAAGAAAGGTGAGTTCAAAATAGACATG 600

DB 541 ACCTGGACATGCACTGTCTTGCAAGAACCAAGAAAGGTGAGTTCAAAATAGACATG 600  
 QY 601 GTGCTAGCTTTCCAGAAAGCTCCGAGGTGCGGTAAGTGGGGAGGCGGTTCAAGCGGA 660  
 DB 601 GTGCTAGCTTTCCAGAAAGCTCCGAGGTGCGGTAAGTGGGGAGGCGGTTCAAGCGGA 660  
 QY 661 GGTGATCCGTTGGCGGAGGAGTCCGGCGGGGGTGAAGCGGGGGTGGCGGCTCCGAGGC 720  
 DB 661 GGTGATCCGTTGGCGGAGGAGTCCGGCGGGGGTGAAGCGGGGGTGGCGGCTCCGAGGC 720  
 QY 721 GGAAGTTCAAGTGAAGTGTCTCGAGTCTGGGGCTGAGAGTGAAGAGCTGGTCTCG 780  
 DB 721 GGAAGTTCAAGTGAAGTGTCTCGAGTCTGGGGCTGAGAGTGAAGAGCTGGTCTCG 780  
 QY 781 GTAAGGTCCTCCGAAAGCTCTGAGAGACACTTTCATAGATTAATGTTTACTGGAGT 840  
 DB 781 GTAAGGTCCTCCGAAAGCTCTGAGAGACACTTTCATAGATTAATGTTTACTGGAGT 840  
 QY 841 CGACAGGCTCTGAGCAAGGCTTGAAGTGAAGAGATCATCATATCTTGATGTA 900  
 DB 841 CGACAGGCTCTGAGCAAGGCTTGAAGTGAAGAGATCATCATATCTTGATGTA 900  
 QY 901 GCACACTTACGACCCGACCTCCAGGCGAAGTCAAGTACCGCGGACAAATCCAGAC 960  
 DB 901 GCACACTTACGACCCGACCTCCAGGCGAAGTCAAGTACCGCGGACAAATCCAGAC 960  
 QY 961 ACAATCTACCTGAGCTGCGGAATCTAAGATCTGACATGACGAGCGGTATATTTCTGTGCG 1020  
 DB 961 ACAATCTACCTGAGCTGCGGAATCTAAGATCTGACATGACGAGCGGTATATTTCTGTGCG 1020  
 QY 1021 GGAATGTACGAGGAGAGCGGACGAGCGAGGGAATATGATTAATGAGTTTCTGAACAT 1080  
 DB 1021 GGAATGTACGAGGAGAGCGGACGAGCGAGGGAATATGATTAATGAGTTTCTGAACAT 1080  
 QY 1081 TGGGGCCAGGGAACCTGTGTCACGTCACCTGAGTGGGGGCTCCGAGGTGGTG 1140  
 DB 1081 TGGGGCCAGGGAACCTGTGTCACGTCACCTGAGTGGGGGCTCCGAGGTGGTG 1140  
 QY 1141 AGCGTGGGGGAGATCTGAACTCGAGTTGAGCGAGTCCAGCCACCTGTCTGTGCT 1200  
 DB 1141 AGCGTGGGGGAGATCTGAACTCGAGTTGAGCGAGTCCAGCCACCTGTCTGTGCT 1200  
 QY 1201 CCAGGGGAAAGACCCCTCTCTGCGAGGCGCAGTGAAGTGTATAGCACTTAAGCC 1260  
 DB 1201 CCAGGGGAAAGACCCCTCTCTGCGAGGCGCAGTGAAGTGTATAGCACTTAAGCC 1260  
 QY 1261 TGGTACACAGAAACCTGGCCAGGCTCCAGGCTCCATATATGGTGCATCCACAGG 1320  
 DB 1261 TGGTACACAGAAACCTGGCCAGGCTCCAGGCTCCATATATGGTGCATCCACAGG 1320  
 QY 1321 GCCACCGGTGCCAGCCAGGTTCAAGTGGCAGTGGGTCTGGGGCAGAAATTCATCTCACC 1380  
 DB 1321 GCCACCGGTGCCAGCCAGGTTCAAGTGGCAGTGGGTCTGGGGCAGAAATTCATCTCACC 1380  
 QY 1381 ATCAGCAGCTGCAAGTCTGAAGATTTTGCAGTTTATTACTGTACAGCATTAATCTGG 1440  
 DB 1381 ATCAGCAGCTGCAAGTCTGAAGATTTTGCAGTTTATTACTGTACAGCATTAATCTGG 1440

# RESULT 2

ADN07735 standard; DNA; 8911 BP.

ADN07735;

15-JUL-2004 (first entry)

Expression vector DNA.

Immunogenic complex; gp120; CD4; HIV; HIV infection; anti-HIV; circular;

cyclic; ds.

OS	Synthetic.
XX	US2004076636-A1.
PN	
XX	
PD	22-APR-2004.
XX	
PF	02-JUL-2003; 2003JUS-00612192.
PR	
XX	
PR	07-MAY-1993; 93US-00060926.
PR	06-MAY-1994; 94WO-US005020.
PR	20-DEC-1995; 95US-00464680.
PR	11-MAY-1998; 98US-00075544.
PR	07-JAN-2000; 2000US-00479675.
PR	17-JUN-2001; 2001US-00905962.
XX	
PA	(PALR/) PAL R.
PA	(MARK/) MARKHAM P.
PA	(KEEN/) KEEN T.
PA	(WHIT/) WHITNEY S.
PA	(KALY/) KALIYANARAMAN V S.
P1	Pal R, Markham P, Keen T, Whitney S, Kaliyanaraman VS;
DR	WPI, 2004-387924/36.
XX	
PT	Immunogenic complex, useful for treating HIV infections, comprises gp120
PT	covalently bonded to CD4 equivalent molecule, fragment of CD4 or its
PT	equivalent.
PS	Disclosure; SEQ ID NO 3; 26pp; English.
CC	The invention relates to an immunogenic complex comprising gp120
CC	covalently bonded to a CD4 equivalent molecule, fragment of CD4 or its
CC	equivalent. The invention also relates to a composition comprising the
CC	complex and a carrier, an antibody reactive with the complex, an
CC	immunolised cell line that produces the complex, a method of detecting
CC	the HIV antigen in a test fluid, involving contacting the test fluid with
CC	an antibody raised against the immunogenic complex and detecting the
CC	presence of immune complexes formed between the antigen in the test fluid
CC	and the antibody, and a vaccine comprising an immunogenically effective
CC	amount of the immunogenic complex. The immunogenic complex is useful for
CC	raising neutralising antibodies against HIV, which involves administering
CC	the complex to a subject in a carrier, and for treating HIV infections.
CC	This sequence represents expression vector DNA used in the scope of the
CC	invention.
SO	Sequence 8911 BP; 2188 A; 2221 C; 2227 G; 2275 T; 0 U; 0 Other;
Query Match	43.4%; Score 625; DB 12; Length 8911;
Best Local Similarity	100.0%; Pred. No. 1.9e-129;
Matches 625; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 ATGAACGGGAGATCGCTTTTGGAGACTTGTCGTGTCGCATCTGGCGTCTCTCCA 60
DB	3815 ATGAACCGGAGATCGCTTTTGGAGACTTGTCGTGTCGCATCTGGCGTCTCTCCA 3874
OY	61 GCAGCACTCAGGGAAGAAGTGTGTGCTGGCAAAAAAGGGATTACAGTGACACC 120
DB	3875 GCAGCACTCAGGGAAGAAGTGTGTGCTGGCAAAAAAGGGATTACAGTGACACC 3933
OY	121 TGTCAGCTTCCACAGAGAAGCATATTCCTCACTGGAAAACTCCAACGATTAAG 180
DB	3935 TGTCAGCTTCCACAGAGAAGCATATTCCTCACTGGAAAACTCCAACGATTAAG 3994
OY	181 ATTCTGGAAAATCAGGGCTCCTTTAATAAGATCCATCAACTGAATATGCGCT 240
DB	3995 ATTCTGGAAAATCAGGGCTCCTTTAATAAGATCCATCAACTGAATATGCGCT 4053
OY	241 GACTCAAGAAGAGCTTTTGGACAAGAAATCTCCCTGATCATCAAGATCTTAAAG 300
DB	4055 GACTCAAGAAGAGCTTTTGGACAAGAAATCTCCCTGATCATCAAGATCTTAAAG 4113
OY	301 ATAGAAGACTCAGATCTTACTTGTGATGAGAGACAGAGAAGAGGATGCAATTG 360

Dd		4115	ATAGAAGCTCAGATACTTACATCTGTGAAGTGGAAGGACCAGAAAGAGAGGTGCATTGG	4174
Oy		361	CTAGTGTTCGAGATTGACTGCCAAGCTCTGACACCACTGCTTTGAGGGGCGAGGCTTAAC	420
Dd		4175	CTAGTGTTCGAGATTGACTGCCAAGCTCTGACACCACTGCTTTGAGGGGCGAGGCTTAAC	4234
Oy		421	CTGACCTTTGGAGAGCCCCCTGGTAGTAGAGCCCTCACTGCAATGTATAGAGTCCAAAGGGGT	480
Dd		4235	CTGACCTTTGGAGAGCCCCCTGGTAGTAGAGCCCTCACTGCAATGTATAGAGTCCAAAGGGGT	4294
Oy		481	AAAAACATACAGGGGGGGAAGAACCTCTCCGTGTCTAGCTGAGAGCTCCAGATAGTGGC	540
Dd		4295	AAAAACATACAGGGGGGGAAGAACCTCTCCGTGTCTAGCTGAGAGCTCCAGATAGTGGC	4354
Oy		541	ACCTGACATGACACTGTCTTTGACAGAACCAAGAGGTGAGATTCAAATATGACATCTGTG	600
Dd		4355	ACCTGACATGACACTGTCTTTGACAGAACCAAGAGGTGAGATTCAAATATGACATCTGTG	4414
Oy		601	GTCGTAGCTTTCCAGAAAGGCTCCG	625
Dd		4415	GTCGTAGCTTTCCAGAAAGGCTCCG	4439

RESULT 3  
AAF82582  
ID AAF82582 standard; cDNA; 1377 BP.  
XX  
XX AAF82582;  
XX  
DT 18-JUN-2001 (first entry)  
XX  
DE Human CD4 gene T4.  
XX  
XX Human; CD4; T4; CD4 fusion protein; oligomerization;  
KW receptor-ligand interaction inhibition; surface plasmon resonance; SPR;  
KW T cell receptor binding; MHC binding; carcinoma; autoimmune disease;  
KW multiple sclerosis; human immunodeficiency virus; HIV; diabetes;  
KW rheumatoid arthritis; immune disorder; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1377  
FT /\*tag= a  
FT /product= "human CD4"  
FT sig\_peptide 1..75  
FT /\*tag= b  
FT mat\_peptide 76..1374  
FT /\*tag= c  
XX  
PN WO200122084-A2.  
XX  
PD 29-MAR-2001.  
XX  
PF 18-SEP-2000; 2000WO-GB003579.  
XX  
PR 21-SEP-1999; 99GB-00022352.  
XX  
PA (AVID-) AVIDEX LTD.  
XX  
PI Jakobsen BK;  
XX  
DR WPI; 2001-273470/28.  
XX P-PDBj; AAB01502.  
PT Sequential screening of candidate compounds library for those which  
PT inhibit binding of low affinity receptor-ligand interaction having fast  
PT binding kinetics, using interfacial optical assay.  
XX  
PS Disclosure; Fig 13; 91pp; English.  
XX  
CC The present sequence encodes human CD4. Human CD4 extracellular domains 1

CC and 2 were used in the construction of CD4 oligomerisation fusion  
 CC proteins. The fusion proteins contain an oligomerisation domain that  
 CC enables the proteins to bind to one another to form oligomers. The  
 CC oligomers may be used in an invention relating to a method for screening  
 CC for compounds with the ability to inhibit a low affinity receptor-ligand  
 CC interaction. The method uses an interfacial optical assay, such as  
 CC surface plasmon resonance (SPR). The method is useful for screening  
 CC candidate compounds for the ability to inhibit interaction between  
 CC MHC/peptide complex and T cell receptor, and MHC/peptide complex and CD8  
 CC or CD4 co-receptor. The compounds identified by the above methods which  
 CC interfere with T cell receptor binding to a particular HLA type molecule  
 CC are useful as immune inhibitors for treating carcinoma, autoimmune  
 CC diseases such as multiple sclerosis, human immunodeficiency virus (HIV)  
 CC infection, rheumatoid arthritis, Hashimoto's disease, insulin dependent  
 CC diabetes, Good pasture's syndrome, uveitis, psoriasis and graft rejection  
 CC  
 XX

SQ Sequence 1377 BP; 351 A; 355 C; 393 G; 278 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 4; Length 1377;  
 Best Local Similarity 99.8%; Pred. No. 1.7e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACCGGGGAGTCCCTTTAGGCACTTGCTGTGCTGCACTGGGCTCTCCCA 60  
 DB 1 ATGACCGGGGAGTCCCTTTAGGCACTTGCTGTGCTGCACTGGGCTCTCCCA 60  
 QY GCAGCCACTCAGGGAAGAAAGTGCTGCTGGCCAAAAGGGATACATGCAATGACC 120  
 DB GCAGCCACTCAGGGAAGAAAGTGCTGCTGGCCAAAAGGGATACATGCAATGACC 120  
 QY 121 TGTACAGCTTCCGAAAGAAAGAGATATTCATCTGGAAGAACTCAACGATTAAG 180  
 DB 121 TGTACAGCTTCCGAAAGAAAGAGATATTCATCTGGAAGAACTCAACGATTAAG 180  
 QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACCTAAAGGTCATCCAGCTGAATGATCGGCT 240  
 DB 181 ATTCTGGGAATCAGGGCTCTTCTTAACCTAAAGGTCATCCAGCTGAATGATCGGCT 240  
 QY 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 300  
 DB 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 300  
 QY 301 ATGAGAACTCAATATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 360  
 DB 301 ATGAGAACTCAATATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 360  
 QY 361 CTAGTGTTCGGATTGACTGCACTGCAACCACTGCTTCAGGGGCAAGGCTGACC 420  
 DB 361 CTAGTGTTCGGATTGACTGCACTGCAACCACTGCTTCAGGGGCAAGGCTGACC 420  
 QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGT 480  
 DB 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGT 480  
 QY 481 AAAAATCTAAGGGGGGGAAGACCTTCTCGTGTCTCACTGGAAGCTCCAGATATGGC 540  
 DB 481 AAAAATCTAAGGGGGGGAAGACCTTCTCGTGTCTCACTGGAAGCTCCAGATATGGC 540  
 QY 541 ACCGTGACATGCACTGTCTTGAAGAACCAAGAAAGGTGAGTTCAAAATAGACATGTC 600  
 DB 541 ACCGTGACATGCACTGTCTTGAAGAACCAAGAAAGGTGAGTTCAAAATAGACATGTC 600  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
 DB 601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626

RESULT 4  
 AA038761  
 XX AA038761 standard; DNA; 1415 BP.  
 AC  
 XX AA038761;

DT 25-MAR-2003 (revised)  
 DT 22-JUL-1993 (first entry)  
 XX sCD4-I2 lysosomal targeting fusion gene.  
 DE  
 XX Soluble CD4; HIV; Human immunodeficiency Virus; envelope; glycoprotein;  
 KM polymerase chain reaction; lysosomal membrane protein; Lamp-2; 8e.  
 XX  
 XX Synthetic.  
 OS  
 FH Key Location/Qualifiers  
 FT misc\_feature 1..6  
 FT /tag= a  
 FT /note= "EcoRI restriction site"  
 FT misc\_feature 1275..1280  
 FT /tag= b  
 FT /note= "XbaI restriction site"  
 FT misc\_feature 1410..1415  
 FT /tag= c  
 FT /note= "SalI restriction site"

W09306216-A1.

PD 01-APR-1993.

PF 22-SEP-1992; 92WO-US008090.

PR 26-SEP-1991; 91US-00766963.

PA (OKLA-) OKLAHOMA MED RES FOUND.

PI Tang JUN, Lin XL;

PI WPI; 1993-117537/14.

PT New fusion protein used in gene therapy for treating AIDS - comprises  
 PT protein which binds to retroviral envelope protein which targets fusion  
 PT protein to lysosome.

PS Claim 12; Page 14-15; 47Pp; English.

CC Primers P-1 and P-2 (AA038748 and AA038749) were used to amplify an EcoRI  
 CC -XbaI fragment of plasmid p14B containing the sCD4 sequence. The PCR  
 CC product was ligated to an XbaI-SalI fragment containing the sequence  
 CC encoding lysosomal membrane protein (LAMP-2) (PCR amplified from a human  
 CC liver lambda gp10 library using primers AA038756 and AA038757). The  
 CC resulting fusion sequence sCD4-I2 can be used to interfere with the  
 CC normal function of HIV and to direct the virus to lysosomes. (Updated on  
 CC 25-MAR-2003 to correct PN field.)  
 CC  
 XX

SQ Sequence 1415 BP; 360 A; 365 C; 393 G; 297 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 2; Length 1415;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACCGGGGAGTCCCTTTAGGCACTTGCTGTGCTGCACTGGGCTCTCCCA 60  
 DB 81 ATGACCGGGGAGTCCCTTTAGGCACTTGCTGTGCTGCACTGGGCTCTCCCA 140  
 QY 61 GCAGCCACTCAGGGAAGAAAGTGCTGCTGGCCAAAAGGGATACATGCAATGACC 120  
 DB 141 GCAGCCACTCAGGGAAGAAAGTGCTGCTGGCCAAAAGGGATACATGCAATGACC 200  
 QY 121 TGTACAGCTTCCGAAAGAAAGAGATATTCATCTGGAAGAACTCAACGATTAAG 180  
 DB 201 TGTACAGCTTCCGAAAGAAAGAGATATTCATCTGGAAGAACTCAACGATTAAG 260  
 QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACCTAAAGTCACTCAAGCTGAATGATCGGCT 240  
 DB 261 ATTCTGGGAATCAGGGCTCTTCTTAACCTAAAGTCACTCAAGCTGAATGATCGGCT 320  
 QY 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATTTAAG 300

Db 321 GACTCAAGAAAGCCTTTGGAGCAAGAACTCCCTGATCATCAAGAACTTAAAG 380  
 Qy 301 ATAGAACTCAGATTAATCATCTGTGAGTGGAGCAAGAGAGAGAGTCAATTG 360  
 Db 381 ATAGAACTCAGATTAATCATCTGTGAGTGGAGCAAGAGAGAGAGTCAATTG 440  
 Qy 361 CTAGTGTGGATTAATCATCTGTGAGTGGAGCAAGAGAGAGTCAATTG 420  
 Db 441 CTAGTGTGGATTAATCATCTGTGAGTGGAGCAAGAGAGAGTCAATTG 500  
 Qy 421 CTAGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 Db 501 CTAGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 560  
 Qy 481 AAAAATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
 Db 561 AAAAATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 620  
 Qy 541 ACCTGACATGACATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
 Db 621 ACCTGACATGACATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 680  
 Qy 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
 Db 681 GTGCTAGCTTTCCAGAAAGGCTCCAG 706  
 RESULT 5  
 ADA44806  
 ID ADA44806 standard; DNA; 1419 BP.  
 AC ADA44806;  
 XX  
 DT 04-DEC-2003 (first entry)  
 DE CD4/TCR CD3epsilon chain chimeric protein CD4epsilon1on15 DNA, SEQ ID NO:1.  
 XX  
 KM HIV-1 infection; human immunodeficiency virus-1; CD4+ cell; chimeric CD4;  
 KM endoplasmic reticulum; ER retention; envelope protein gp160;  
 KM T cell receptor CD3epsilon chain; C-terminal domain; CD4epsilon1on15;  
 KM gene therapy; human; gene; ds.  
 XX  
 OS Chimeric.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1419  
 FT /\*tag= a  
 FT /product= "CD4epsilon1on15"  
 FT /note= "No stop codon given"  
 FT  
 PN WO2003076468-A1.  
 PD 18-SEP-2003.  
 XX  
 PF 14-MAR-2003; 2003WO-ES000120.  
 XX  
 PR 14-MAR-2002; 2002ES-0000616.  
 XX  
 PA (NNSU) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
 PI Alarcon Sanchez BJ, San Jose Martinez ME, Zaldivar Notario I;  
 XX Gomez Buendia M;  
 DR WPI; 2003-779059/73.  
 DR P-PSDB; ADA44807.  
 XX  
 PT Composition for treating or preventing human immune deficiency virus,  
 PT comprises CD4 chimeric protein having a protective effect in trans, or  
 PT related nucleic acid.  
 XX

PS Claim 5; Page 31-33; 43pp; Spanish.  
 XX  
 CC The invention relates to a composition for the treatment or prevention of  
 CC human immunodeficiency virus-1 (HIV-1) infection. The composition  
 CC comprises CD4+ cells that have been transduced with a vector that encodes  
 CC a chimeric CD4 molecule which is capable of being retained in the  
 CC endoplasmic reticulum (ER). The invention also encompasses the use of a  
 CC soluble protein factor produced by CD4+ cells that have been transduced  
 CC with a vector encoding a chimeric CD4 protein, and the use of an  
 CC expression system encoding a chimeric CD4 protein. The ER-localised  
 CC chimeric CD4 molecule binds to the HIV-1 envelope protein gp160,  
 CC resulting in HIV-1 retention in the ER and thereby preventing viral  
 CC replication. In a specific embodiment, the chimeric CD4 molecule  
 CC comprises CD4 fused to 15 amino acids of the C-terminal domain of the T  
 CC cell receptor CD3epsilon chain; this chimeric CD4 molecule is designated  
 CC CD4epsilon1on15 (ADA44807). A known chimeric CD4 of similar structure but  
 CC containing only 10 amino acids from CD3epsilon can also be used.  
 CC Compositions of the invention have an in trans effect on the replication  
 CC of HIV-1, and may be used to treat and prevent HIV-1 infection. The  
 CC present sequence represents a DNA encoding the chimeric CD4 molecule  
 CC CD4epsilon1on15, which is specifically claimed for use in compositions of  
 CC the invention.  
 XX  
 SQ Sequence 1419 BP; 362 A; 367 C; 405 G; 285 T; 0 U; 0 Other;  
 Query Match 43.4%; Score 624.4; DB 10; Length 1419;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 ATGAAACCGGGAGTCCCTTTTAAAGCACTTGTGAGTGTGCACTGCGCTCTCCCA 60  
 Db 1 ATGAAACCGGGAGTCCCTTTTAAAGCACTTGTGAGTGTGCACTGCGCTCTCCCA 60  
 Qy 61 GCAGCCACTCAGGAGAAAGAGTGTGCTGGGCAAAAAAGGGATACAGTGAAGTACC 120  
 Db 61 GCAGCCACTCAGGAGAAAGAGTGTGCTGGGCAAAAAAGGGATACAGTGAAGTACC 120  
 Qy 121 TGTACAGCTTCCAGAAAGAGCATATCAATTCATGGAATACTCCAACTAGATTAAG 180  
 Db 121 TGTACAGCTTCCAGAAAGAGCATATTCATGGAATACTCCAACTAGATTAAG 180  
 Qy 181 ATTCTGGAAATCAGAGGCTCTTCTTAATCAAGTTCATCAAGTGAATGCGGCT 240  
 Db 181 ATTCTGGAAATCAGAGGCTCTTCTTAATCAAGTTCATCAAGTGAATGCGGCT 240  
 Qy 241 GACTCAAGAAAGAGCTTTGGAGCAAGAACTTCCCTGTATCATCAAGATCTTAAG 300  
 Db 241 GACTCAAGAAAGAGCTTTGGAGCAAGAACTTCCCTGTATCATCAAGATCTTAAG 300  
 Qy 301 ATGAAAGACTCAGATTAATCATCTGTGAGAGTGAAGAGAGAGAGAGAGAGAG 360  
 Db 301 ATGAAAGACTCAGATTAATCATCTGTGAGAGTGAAGAGAGAGAGAGAGAGAG 360  
 Qy 361 CTAGTGTGGATTAATCATCTGTGAGAGTGAAGAGAGAGAGAGAGAGAGAGAG 420  
 Db 361 CTAGTGTGGATTAATCATCTGTGAGAGTGAAGAGAGAGAGAGAGAGAGAGAG 420  
 Qy 421 CTAGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 Db 421 CTAGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 Qy 481 AAAAATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
 Db 481 AAAAATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
 Qy 541 ACCTGACATGACATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
 Db 541 ACCTGACATGACATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
 Qy 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
 Db 601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626

XX	Sequence 1421 BP; 355 A; 382 C; 401 G; 283 T; 0 U; 0 Other;
XX	
CC	normal function of HIV and to direct the virus to lysosomes. (Updated on
CC	25-MAR-2003 to correct PN field.)
CC	
CC	liver lambda gpl0 library using primers AAQ38754 and AAQ38755). The
CC	encoding lysosomal membrane protein (LAMP-1) (PCR amplified from a human
CC	-XbaI fragment of plasmid pT4B containing the SCD4 sequence. The PCR
CC	product was ligated to an XbaI-SalI fragment containing the sequence
CC	primers P-1 and P-2 (AAQ38748 and AAQ38749) were used to amplify an EcoRI
XX	
PS	Claim 12; Page 13-14; 47pp; English.
PT	New fusion protein used in gene therapy for treating AIDS - comprises
PT	protein which binds to retroviral envelope protein which targets fusion
PT	protein to lysosome.
XX	
DR	WPI; 1993-117537/14.
XX	
P1	Tang JUN, Lin XL;
XX	
PA	(OKLA-) OKLAHOMA MED RES FOUND.
XX	
PR	26-SEP-1991; 91US-00766963.
XX	
PF	22-SEP-1992; 92WO-US008090.
XX	
PD	01-APR-1993.
XX	
PN	WO9306216-A1.
XX	
FT	misc_feature
FT	/tag= c
FT	/note= "SalI restriction site"
FT	1416..1421
FT	/tag= b
FT	/note= "XbaI restriction site"
FT	1275..1280
FT	/tag= a
FT	/note= "EcoRI restriction site"
FT	misc_feature
FT	Key
FT	Location/Qualifiers
FT	1..6
FT	misc_feature
AC	AAQ38760; standard; DNA; 1421 BP.
ID	AAQ38760
RESULT 6	

Query Match	Similarity	43.4%	Score 624.4	DB 2	Length 1421
Best Local	Similarity	99.8%	Pred. No. 1.8e-129		
Matches	Conservative	0	Mismatches 1	Indels 0	Gaps 0
QY	1	ATGAACCGGGAGATCCCTTTTAGGCACTTGTCTTGCTGCAACTGGCGTCTCTCCCA	60		
Db	81	ATGAACCGGGAGATCCCTTTTAGGCACTTGTCTTGCTGCAACTGGCGTCTCTCCCA	140		
QY	61	GCAGGCACTCAGGGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGAACCTGAC	120		
Db	141	GCAGGCACTCAGGGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGAACCTGAC	200		
QY	121	TGTACAGTTCCCAAGAAAGCATACATTCCACTGGAAAACTCCACCAATGAATTAAG	180		

Db	201	TGTACAGCTTCCAGAAAGAGCATACATTCACCTGGAAAAAACTCCACACGATTAAG	260
OY	181	ATTCTGGGAAATCAGGGCTCTTCTTAACTAAAGTCCATCCAGCTGAATGATCGCGCT	240
Db	261	ATTCTGGGAAATCAGGGCTCTTCTTAACTAAAGTCCATCCAGCTGAATGATCGCGCT	320
OY	241	GACTCAAGAAGAGCTTTGGGACCAAGAAACTTCCCGCTGATCATCAAGAACTTTAAG	300
Db	321	GACTCAAGAAGAGCTTTGGGACCAAGAAACTTCCCGCTGATCATCAAGAACTTTAAG	380
OY	301	ATAGAAAGCTCAGATATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCATTG	360
Db	381	ATAGAAAGCTCAGATATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCATTG	440
OY	361	CTAGTGTGGGATTGTGCTGCAACTGTCAACCCACCTGCTTCAGGGGACAGCTGACC	420
Db	441	CTAGTGTGGGATTGTGCTGCAACTGTCAACCCACCTGCTTCAGGGGACAGCTGACC	500
OY	421	CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCACTGATGACATGTAGAGTCCAAAGGGT	480
Db	501	CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCACTGATGACATGTAGAGTCCAAAGGGT	560
OY	481	AAAAAATATCAGAGGGGGGAAAGACCTTCACCGTGTCTACGTGGAGCTCCAGGATAGTGGC	540
Db	561	AAAAAATATCAGAGGGGGGAAAGACCTTCACCGTGTCTACGTGGAGCTCCAGGATAGTGGC	620
OY	541	ACCTGGACATGCACTGTCTTTGACAAACCAAGAAAGGTGAGTTCAAAATATGACATCTGT	600
Db	621	ACCTGGACATGCACTGTCTTTGACAAACCAAGAAAGGTGAGTTCAAAATATGACATCTGT	680
OY	601	GTCCTAGCTTTCCAGAAAGGCTCCGG	626
Db	681	GTCCTAGCTTTCCAGAAAGGCTCCAG	706

```

RESULT 7
AAQ38759
ID    AAQ38759 standard; DNA; 1448 BP.
XX
XX    AAQ38759;
AC
XX
XX    25-MAR-2003 (revised)
DT
XX    22-JUL-1993 (first entry)
XX
XX    sCD4-HAP lysosomal targeting fusion gene.
DE
XX
XX    Soluble CD4; HIV; Human immunodeficiency virus; envelope; ss;
KW    glycoprotein; polymerase chain reaction; human acid phosphatase.
XX
XX    Synthetic.
OS
XX
XX    Key
FH    Location/Qualifiers
FH    misc_feature
FT    1..6
FT    /*tag= a
FT    /note= "Scor1 restriction site"
FT    misc_feature
FT    144..1448
FT    /*tag= b
FT    /note= "Sall restriction site"
XX
XX    WO9306216-A1.
XX
XX    01-APR-1993.
XX
XX    22-SEP-1992; 92WO-US008090.
XX
XX    26-SEP-1991; 91US-00766963.
XX
XX    (OKLA-) OKLAHOMA MED RES FOUND.
XX
XX    Tang JTN, Lin XL;
XX
XX    WPI, 1993-117537/14.
XX

```



PT New fusion protein used in gene therapy for treating AIDS - comprises  
PT protein which binds to retroviral envelope protein which targets fusion  
PT protein to lysosome.

PS Claim 12; Page 12-13; 47pp; English.

CC Primers P-1 and P-2 (AAQ38748 and AAQ38749) were used to amplify an EcoRI  
CC -XbaI fragment of plasmid pT48 containing the SCD4 sequence. The PCR  
CC product was ligated to an XbaI-SalI fragment containing the sequence  
CC encoding human acid phosphatase (HAP) lysosomal targeting protein (PCR  
CC amplified from a human liver lambda gp10 library using primers AAQ38752  
CC and AAQ38753). The resulting fusion sequence SCD4-HAP can be used to  
CC interfere with the normal function of HIV and to direct the virus to  
CC lysosomes. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 1448 BP; 357 A; 395 C; 404 G; 292 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 2; Length 1448;  
Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTGTGCTGCAACTGGCGCTCTCCCA 60  
DB 81 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTGTGCTGCAACTGGCGCTCTCCCA 140  
QY 61 GCAGCCACTCAGGAGAAAGAGTGTGTGCTGGCAAAAAAGGGAGATACAGTGAATGACC 120  
DB 141 GCAGCCACTCAGGAGAAAGAGTGTGTGCTGGCAAAAAAGGGAGATACAGTGAATGACC 200  
QY 121 TGTAAGCTTCCGAAAGAGAGATCAATTCCTGAAATACTCCAAACAGATAAG 180  
DB 201 TGTAAGCTTCCGAAAGAGAGATCAATTCCTGAAATACTCCAAACAGATAAG 260  
QY 181 ATTCTGGGAATCGGGCTCTTCTTAAGTAAAGTTCATTCAGCTGAATGATCGGCT 240  
DB 261 ATTCTGGGAATCGGGCTCTTCTTAAGTAAAGTTCATTCAGCTGAATGATCGGCT 320  
QY 241 GACTCAAGAGAGAGCTTTGGGACCAAGAACTCCCGCTGATCATCAAGATCTTAAG 300  
DB 321 GACTCAAGAGAGAGCTTTGGGACCAAGAACTCCCGCTGATCATCAAGATCTTAAG 380  
QY 301 ATGAAGAGCTCAGATTAATCATCTGTGAAGTGAAGCAAGAGAGAGTGAATG 360  
DB 381 ATGAAGAGCTCAGATTAATCATCTGTGAAGTGAAGCAAGAGAGAGTGAATG 440  
QY 361 CTAGTGTTCGATTAATCATCTGTGAAGTGAAGCAAGAGAGAGTGAATG 420  
DB 441 CTAGTGTTCGATTAATCATCTGTGAAGTGAAGCAAGAGAGAGTGAATG 500  
QY 421 CTGACCTTGGAGAGCCCGCTGTAGTAGCCCTCAGTGAATGAGAGTCAAGGGGT 480  
DB 501 CTGACCTTGGAGAGCCCGCTGTAGTAGCCCTCAGTGAATGAGAGTCAAGGGGT 560  
QY 481 AAAAATACAGAGGGGGGAAAGCCCTCTCGTGTCTCAGCTGAGCTCCAGATATGAGC 540  
DB 561 AAAAATACAGAGGGGGGAAAGCCCTCTCGTGTCTCAGCTGAGCTCCAGATATGAGC 620  
QY 541 ACCTGCAATGCACTGTCTTGCAGAACCAAGAAAGGTGAGTTCAAAATATGACATGTG 600  
DB 621 ACCTGCAATGCACTGTCTTGCAGAACCAAGAAAGGTGAGTTCAAAATATGACATGTG 680  
QY 601 GTGCTAGCTTCCGAGAGGCTCCGG 626  
DB 681 GTGCTAGCTTCCGAGAGGCTCCAG 706

RESULT 8  
ID AAAS0662  
AC AAAS0662; standard; DNA; 1714 BP.

XX 09-JAN-2001 (first entry)

XX DNA encoding CD4-IgM fusion protein CH4Mmu.  
DE CD4; IgM; human; CD4Mmu; fusion protein; immunoglobulin; HIV; SIV; gp120;  
KW therapy; diagnosis; de.  
KM  
XX Homo sapiens.

Key	Location/Qualifiers
FT exon	1..1308
FT	/*tag= b
FT	/number= 1
FT CDS	111..1714
FT	/*tag= a
FT	/partial
FT	/note= "contains an intron"
FT intron	1309..1664
FT	/*tag= c
FT	/number= 1
FT exon	1665..1714
FT	/*tag= d
FT	/partial
FT	/number= 2

XX US6117656-A.  
XX 12-SEP-2000.  
XX 07-JUN-1995; 95US-00479353.  
XX 22-JAN-1988; 88US-00147351.  
XX 23-JAN-1989; 89US-00295956.  
XX 09-JUN-1992; 92US-00896781.  
XX 12-APR-1993; 93US-00057952.  
XX 04-FEB-1994; 94US-00191708.

PA (GENO) GEN HOSPITAL CORP.

PI Seed B;

XX WPI; 2000-586558/55.

DR P-PSDB; AAB19509.

PT CD4-immunoglobulin fusion proteins, useful for targeting gp120 of HIV or  
PT SIV.

XX Example 1; Col 41-50; 39pp; English.

XX The present sequence is that of DNA encoding fusion protein CD4Mmu (see  
XX AAB19509) comprising the extracellular portion of CD4, which binds to HIV  
XX gp120, linked at its C-terminus to a human IgM heavy chain polypeptide.  
XX To obtain the construct, DNA encoding CD4 was linked to IgM DNA at the  
XX Met2 site upstream of the CH1 region. A plasmid containing this genetic  
XX construct is deposited in Escherichia coli MC1061/P3 as ATCC 67609.  
XX Fusion protein CD4Mmu and a nucleic acid encoding it are claimed. Also  
XX claimed are a vector comprising the nucleic acid, and a method of  
XX producing the fusion protein in secreted form using a transformed host  
XX cell. The fusion protein may further comprise a therapeutic agent,  
XX radiolabel or NMR imaging agent. The fusion protein can be administered  
XX to an animal (including humans) for treatment of HIV or SIV infection.  
XX and can also be used in assays for HIV or SIV, imaging and tissue stains.  
XX IgM fusion proteins such as CD4Mmu provide complement-mediated immunity

XX Sequence 1714 BP; 411 A; 503 C; 467 G; 333 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 3; Length 1714;  
Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTGTGCTGCAACTGGCGCTCTCCCA 60  
DB 111 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTGTGCTGCAACTGGCGCTCTCCCA 170

```

QY      61 GCAGCCACTCAGGAGAAAGTGTGCTGTGGCAAAAAGGGGATACAGTGAACCTGACC 120
      171 GCAGCCACTCAGGAGAAAGTGTGCTGTGGCAAAAAGGGGATACAGTGAACCTGACC 230
QY      121 TGTACAGCTTCCAGAAAGAGATACATTCACACTGGAAAAAATCCAAACCAATTAAG 180
      231 TGTACAGCTTCCAGAAAGAGATACATTCACACTGGAAAAAATCCAAACCAATTAAG 290
QY      181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCAAGCTGAATGATCGGCT 240
      291 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCAAGCTGAATGATCGGCT 350
QY      241 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 300
      351 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 410
QY      301 ATGAGAACTCAATTAATTAATCTGTGAAGTGAAGACCAAGAAAGAGGTGCAATTG 360
      411 ATGAGAACTCAATTAATTAATCTGTGAAGTGAAGACCAAGAAAGAGGTGCAATTG 470
QY      471 CTAGTGTTCGGATTGATGCTGCAACCTGCAACCTGCTTCAGGGGCAAGCTGACC 530
      531 CTAGTGTTCGGATTGATGCTGCAACCTGCAACCTGCTTCAGGGGCAAGCTGACC 590
QY      531 CTAGTGTTCGGATTGATGCTGCAACCTGCAACCTGCTTCAGGGGCAAGCTGACC 590
      591 AAAAACAATACAGGGGGGAAAGACCTCTCCGTCTCAAGCTGAGCTCCAGATAGTGGC 650
QY      591 AAAAACAATACAGGGGGGAAAGACCTCTCCGTCTCAAGCTGAGCTCCAGATAGTGGC 650
      651 ACCTGACATGCACTGTCTTGCAGAAACCAAGAAAGGTGAGTTCAAAAATGACATCGTG 710
QY      651 ACCTGACATGCACTGTCTTGCAGAAACCAAGAAAGGTGAGTTCAAAAATGACATCGTG 710
      711 GTGCTAGCTTTCCAGAAAGGCTCCAG 736
QY      711 GTGCTAGCTTTCCAGAAAGGCTCCAG 736

```

## RESULT 9

AAZ44063 ID AAZ44063 standard; DNA; 1714 BP.

XX AAZ44063;

XX 23-MAR-2000 (first entry)

XX Human fusion protein CD4Mg encoding DNA.

XX Fusion protein; human; CD4; IGM; immunoglobulin; gp120;

XX anti-human immunodeficiency virus; CD4Mg; ds.

XX Homo sapiens.

XX Synthetic.

XX US6004781-A.

XX 21-DEC-1999.

XX 04-FEB-1994; 94US-00191708.

XX 23-JAN-1988; 88US-00147351.

XX 23-JAN-1989; 89US-00299596.

XX 09-JUN-1992; 92US-00896781.

XX 12-APR-1993; 93US-00057952.

XX (GEHO ) GEN HOSPITAL CORP.

XX Seed B;

XX WPI; 2000-085792/07.

```

DR      P-PSDB; AAY51080.
PT      Fusion protein useful for the treatment of human immunodeficiency virus.
XX      Example 1; Col 41-50; 39pp; English.
XX      This invention describes a novel nucleic acid (i) encoding a fusion
XX      protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)
XX      and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light
XX      chain (III). The products of the invention have anti-human
XX      immunodeficiency virus (HIV) activity and are capable of binding to
XX      gp120. The fusion protein is useful for treating human immunodeficiency
XX      virus (HIV) or simian immunodeficiency virus (SIV). This sequence encodes
XX      the fusion protein CD4Mg which is constructed from CD4 linked to human
XX      IGM upstream of the CH1 region
SQ      Sequence 1714 BP; 411 A; 503 C; 467 G; 333 T; 0 U; 0 Other;

```

Query Match 43.4%; Score 624.4; DB 3; Length 1714;

Best Local Similarity 99.8%; Pred. No. 1.8e-129;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY      1 ATGAACCGGGAGTCCCTTTAGGCACTGTGCTGTGCTGCAACTGGCGCTCCCA 60
      111 ATGAACCGGGAGTCCCTTTAGGCACTGTGCTGTGCTGCAACTGGCGCTCCCA 170
QY      61 GCAGCCACTCAGGAGAAAGTGTGCTGTGGCAAAAAGGGGATACAGTGAACCTGACC 120
      171 GCAGCCACTCAGGAGAAAGTGTGCTGTGGCAAAAAGGGGATACAGTGAACCTGACC 230
QY      171 GCAGCCACTCAGGAGAAAGTGTGCTGTGGCAAAAAGGGGATACAGTGAACCTGACC 230
      231 TGTACAGCTTCCAGAAAGAGATACATTCACACTGGAAAAAATCCAAACCAATTAAG 180
QY      231 TGTACAGCTTCCAGAAAGAGATACATTCACACTGGAAAAAATCCAAACCAATTAAG 180
      291 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCAAGCTGAATGATCGGCT 240
QY      291 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCAAGCTGAATGATCGGCT 240
      351 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 300
QY      351 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 300
      411 ATGAGAACTCAATTAATTAATCTGTGAAGTGAAGACCAAGAAAGAGGTGCAATTG 360
QY      411 ATGAGAACTCAATTAATTAATCTGTGAAGTGAAGACCAAGAAAGAGGTGCAATTG 360
      471 CTAGTGTTCGGATTGATGCTGCAACCTGCAACCTGCTTCAGGGGCAAGCTGACC 530
QY      471 CTAGTGTTCGGATTGATGCTGCAACCTGCAACCTGCTTCAGGGGCAAGCTGACC 530
      531 CTAGTGTTCGGATTGATGCTGCAACCTGCAACCTGCTTCAGGGGCAAGCTGACC 590
QY      531 CTAGTGTTCGGATTGATGCTGCAACCTGCAACCTGCTTCAGGGGCAAGCTGACC 590
      651 ACCTGACATGCACTGTCTTGCAGAAACCAAGAAAGGTGAGTTCAAAAATGACATCGTG 710
QY      651 ACCTGACATGCACTGTCTTGCAGAAACCAAGAAAGGTGAGTTCAAAAATGACATCGTG 710
      711 GTGCTAGCTTTCCAGAAAGGCTCCAG 736

```

## RESULT 10

AAZ48203 ID AAZ48203 standard; DNA; 1714 BP.

XX AAZ48203;

DT 14-MAR-2000 (first entry)  
 XX DNA sequence encoding CD4-Ig fusion protein CD4Mmu.  
 DE  
 XX HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;  
 KM secreted protein; SIV infection; medication; ds.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 PN CA1340741-C.  
 XX  
 PD 14-SEP-1999.  
 XX  
 PF 20-JAN-1989; 89CA-00588749.  
 XX  
 PR 20-JAN-1989; 89CA-00588749.  
 XX  
 PA (GEHO ) GEN HOSPITAL CORP.  
 XX  
 PI Seed B;  
 XX  
 DR WPI; 2000-063015/06.  
 DR P-PSDB; AAY59170.  
 XX  
 XX  
 PT New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in  
 PT the treatment of HIV or simian immunodeficiency virus infections.  
 XX  
 XX Example 1; Page 47-53; 89pp; English.  
 XX  
 CC The invention provides a fusion gene encoding a fusion protein that  
 CC comprises an extracellular CD4 DNA sequence or its fragment which binds  
 CC to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA  
 CC sequence of an Ig heavy or light chain, where the DNA sequence encoding  
 CC the variable region has been replaced with the DNA sequence which encodes  
 CC extracellular CD4 or its gp120 binding fragment. The fusion protein is  
 CC capable of being secreted. The fusion proteins are useful for treating  
 CC HIV or SIV infections in animals, preferably humans. They are also useful  
 CC for producing medicaments which can be used for treating HIV or SIV  
 CC infections in humans. The present sequence represents the DNA which  
 CC encodes the fusion protein CD4Mmu where the CD4 is linked to human IgG1  
 CC at the Mst2 site upstream of the CH1 region  
 XX  
 XX Sequence 1714 BP; 411 A; 503 C; 467 G; 333 T; 0 U; 0 Other:  
 SQ  
 Query Match 43.4%; Score 624.4; DB 3; Length 1714;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGAACCGGGAGTCCCTTTAGGCACTGCTGTGGTGGCAACTGGGGCTCTCCCA 60  
 DB 111 ATGAACCGGGAGTCCCTTTAGGCACTGCTGTGGTGGCAACTGGGGCTCTCCCA 170  
 QY 61 GCAGCCACTCAGGGAAGAAAGTGTGTCTGGGCAAAAAAGGGGATACAGTGAACCTGACC 120  
 DB 171 GCAGCCACTCAGGGAAGAAAGTGTGTCTGGGCAAAAAAGGGGATACAGTGAACCTGACC 230  
 QY 121 TGTACAGCTTCCCGAAGAGAGATACATTCACCTGGAAAACTCCAACCCAGATTAAG 180  
 DB 231 TGTACAGCTTCCCGAAGAGAGATACATTCACCTGGAAAACTCCAACCCAGATTAAG 290  
 QY 181 ATTCTGGGAATCAGGGGCTCTCTTAACCTAAGGTCATCAAGCTGAATGATGGGCT 240  
 DB 291 ATTCTGGGAATCAGGGGCTCTCTTAACCTAAGGTCATCAAGCTGAATGATGGGCT 350  
 QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
 DB 351 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 410  
 QY 301 ATAGAAGACTCAGTACTTACATCTGTGAAGTGAAGACCAAGAGAGAGAGTGCATTG 360  
 DB 411 ATAGAAGACTCAGTACTTACATCTGTGAAGTGAAGACCAAGAGAGAGAGTGCATTG 470

QY 361 CTAGTGTGGATTTGACTGCACTGCAACCCACTGCTTACGGGGCAGAGCTTGACC 420  
 DB 471 CTAGTGTGGATTTGACTGCACTGCAACCCACTGCTTACGGGGCAGAGCTTGACC 530  
 QY 421 CTGACCTTGGAGAGCCCCCTGTGTAAGCCCTTCAGTCAAGTGAAGTCCAAAGGGGT 480  
 DB 531 CTGACCTTGGAGAGCCCCCTGTGTAAGCCCTTCAGTCAAGTGAAGTCCAAAGGGGT 590  
 QY 481 AAAACATACAGGGGGGGAAGAACCTCTCGTGTCTCAGCTGGAGCTCCAGATTAAGGC 540  
 DB 591 AAAACATACAGGGGGGGAAGAACCTCTCGTGTCTCAGCTGGAGCTCCAGATTAAGGC 650  
 QY 541 ACCTGACATGCACTGTCTTGCAGAAACCAAGAAAGTGAAGTTCAAAATAGACATGCTG 600  
 DB 651 ACCTGACATGCACTGTCTTGCAGAAACCAAGAAAGTGAAGTTCAAAATAGACATGCTG 710  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
 DB 711 GTGCTAGCTTTCCAGAAAGGCTCCAG 736  
 RESULT 11  
 AAA35205  
 ID AAA35205 standard; DNA; 1742 BP.  
 XX  
 AC AAA35205;  
 XX  
 DT 28-JUL-2000 (first entry)  
 XX  
 XX Human adenosine receptor related polynucleotide 2nd SEQ ID NO:79.  
 DE  
 XX Human; adenosine receptor; low adenosine antisense oligonucleotide;  
 KM phosphorothioate; impaired respiration; inflammation; allergy;  
 KM allergic disease; bronchoconstriction; inhibitor; antiinflammatory;  
 KM antiallergic; antiasthmatic; cyostatic; analgesic; impaired airway;  
 KM lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;  
 KM respiratory distress syndrome; pain; cystic fibrosis; emphysema;  
 KM pulmonary hypertension; chronic obstructive pulmonary disease; COPD;  
 KM cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200009525-A2.  
 PN  
 XX 24-FEB-2000.  
 PD  
 XX 03-AUG-1999; 99WO-US017712.  
 XX  
 XX 03-AUG-1998; 98US-0095212P.  
 PR  
 XX (UYEC-) UNIV EAST CAROLINA.  
 PA  
 XX Nyce JW;  
 PI  
 XX WPI; 2000-205971/18.  
 DR  
 XX  
 XX New antisense oligonucleotides useful for treating e.g. pulmonary  
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,  
 PT bronchitis, emphysema, respiratory distress syndrome, ischaemia or  
 PT cancers.  
 XX  
 PS Disclosure; Page 1245; 1343pp; English.  
 XX  
 CC The present invention describes a new composition comprising an antisense  
 CC oligonucleotide (ON) with low adenosine (up to 15%), which targets  
 CC nucleic acids involved in bronchoconstriction, allergies, and/or  
 CC inflammation. The ON can have antiinflammatory, antiallergic,  
 CC antiasthmatic, cyostatic and analgesic activities. The compositions are  
 CC useful for the treatment of diseases associated with inflammation,  
 CC impaired airways, including lung disease and diseases whose secondary  
 CC effects afflict the lungs of a subject. They can be used for treating  
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,  
 CC impaired respiration, respiratory distress syndrome, pain, cystic

CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive  
 CC pulmonary disease (COPD), and cancers such as leukemias, lymphomas,  
 CC carcinomas, and cancers which may metastasize to the lungs, including  
 CC breast and prostate cancer. The reduction of the adenosine content of the  
 CC ONs reduces side effects. The A-containing ONs break down with the  
 CC release of deoxyadenosine which activates adenosine receptors causing  
 CC bronchoconstriction and inflammation. AAA3233 to AAA3512 represent the  
 CC nucleotide sequences given in the sequence listing from the present  
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185  
 CC sequences are also called SEQ ID NO:1 to 185, but the sequences differ  
 CC from the previously named sequences. SEQ ID NO:11 to 1680 (AAA3233 to  
 CC AAA33992) are specifically claimed ONs from the present invention. N.B.  
 CC Sequences given in the disclosure of the present invention do not match  
 CC up with their corresponding SEQ ID NO: sequences given in the sequence  
 CC listing  
 CC XX  
 SQ Sequence 1742 BP; 405 A; 492 C; 466 G; 379 T; 0 U; 0 Other;  
 Query Match 43.4%; Score 624.4; DB 3; Length 1742;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGACCGGGGAGTCCCTTTAGGCACTTGCTGTGCTGCAACTGGGCTCTCCCA 60  
 Db 76 ATGACCGGGGAGTCCCTTTAGGCACTTGCTGTGCTGCAACTGGGCTCTCCCA 135  
 QY 61 GCAGCCACTCAGGAGAAAGTGTCTGTGCTGCAAAAGGGGATACATGCAATGACC 120  
 Db 136 GCAGCCACTCAGGAGAAAGTGTCTGTGCTGCAAAAGGGGATACATGCAATGACC 195  
 QY 121 TGTACAGCTTCCGAAAGAAAGAGATCAATTCACCTGCAAAATCTCAACCGATTAAG 180  
 Db 196 TGTACAGCTTCCGAAAGAAAGAGATCAATTCACCTGCAAAATCTCAACCGATTAAG 255  
 QY 181 ATTCGGGAAATCAGGGCTCTCTTAACTAAAGGTCATCCAGCTGAATGATCGGCT 240  
 Db 256 ATTCGGGAAATCAGGGCTCTCTTAACTAAAGGTCATCCAGCTGAATGATCGGCT 315  
 QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGAAATCTTAAG 300  
 Db 316 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGAAATCTTAAG 375  
 QY 301 ATGAGAACTCAATTAATTAATTTGGAAGTGAAGACCAAGAGGAGGTGCAATTG 360  
 Db 376 ATGAGAACTCAATTAATTAATTTGGAAGTGAAGACCAAGAGGAGGTGCAATTG 435  
 QY 361 CTAGTGTTCGATTAATTAATTTGGAAGTGAAGACCAAGAGGAGGTGCAATTG 420  
 Db 436 CTAGTGTTCGATTAATTAATTTGGAAGTGAAGACCAAGAGGAGGTGCAATTG 495  
 QY 421 CTGACCTTGAAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGAT 480  
 Db 496 CTGACCTTGAAGAGCCCTCTGTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGAT 555  
 QY 481 AAAAACAATAAGGGGGGAAAGACCTCTCTGTGTCTCACTGCAAGTCCAGATATGAGC 540  
 Db 556 AAAAACAATAAGGGGGGAAAGACCTCTCTGTGTCTCACTGCAAGTCCAGATATGAGC 615  
 QY 541 ACCGTGATGACCTGTCTTGAAGAACCAAGAAAGGTGAGTCAAAATAGACATGTG 600  
 Db 616 ACCGTGATGACCTGTCTTGAAGAACCAAGAAAGGTGAGTCAAAATAGACATGTG 675  
 QY 601 GTGTAGCTTTCCAGAAAGGCTCCGG 626  
 Db 676 GTGTAGCTTTCCAGAAAGGCTCCAG 701

RESULT 12  
 AAF21327  
 ID AAF21327 standard; DNA; 1742 BP.  
 AC AAF21327;  
 XX

DT 14-MAR-2001 (first entry)  
 XX Human low adenosine antisense oligonucleotide related sequence #2894.  
 XX  
 CC Low adenosine antisense oligonucleotide; phosphorothioate; allergy;  
 CC human; airway disorder; bronchoconstriction; lung inflammation;  
 CC surfactant depletion; respiratory; bronchodilator; antiinflammatory;  
 CC immunosuppressive; antiallergic; hypotensive; cytostatic;  
 CC respiratory obstruction; pulmonary obstruction; impeded respiration;  
 CC surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;  
 CC respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;  
 CC pulmonary hypertension; emphysema; pulmonary transplantation rejection;  
 CC chronic obstructive pulmonary disease; pulmonary infection; bronchitis;  
 CC cancer; ss.  
 CC XX  
 OS Homo sapiens.  
 XX  
 PN W0200062736-A2.  
 PD 26-OCT-2000.  
 PF 24-MAR-2000; 2000WO-US008020.  
 PR 06-APR-1999; 99US-0127958P.  
 XX  
 PA (UYEC-) UNIV EAST CAROLINA.  
 XX  
 PI (NYCE-) NYCE J W.  
 XX  
 PI Nyce JW;  
 XX  
 DR WPI; 2000-679539/66.  
 XX  
 PT Low adenosine (A) content antisense oligonucleotides which do not trigger  
 PT adenosine receptors during metabolism, useful e.g. for treating cancers  
 PT and respiratory obstructions.  
 XX  
 PS Disclosure, Page 1329; 1592pp; English.  
 XX  
 CC The present invention describes low adenosine (A) content antisense  
 CC oligonucleotides and compositions (I) comprising them. In the antisense  
 CC oligonucleotides the A is replaced by a 'Universal' or alternative base.  
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,  
 CC immunosuppressive, antiallergic, hypotensive and cytostatic activities.  
 CC The antisense oligonucleotides and (I) can be used to down-regulate the  
 CC expression and/or activity of target polypeptides associated with  
 CC lung/respiratory disorders and malignancies, such as stimulating and  
 CC activating peptide factors and transmitters, transcription factors,  
 CC immunoglobulin and antibodies, antibody receptors, cytokines and  
 CC chemokines, endogenously produced specific and non-specific enzymes,  
 CC binding proteins, adhesion molecules and their receptors, cytokine and  
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central  
 CC nervous system (CNS) and peripheral nervous and non-nervous system  
 CC receptors, CNS and peripheral nervous and non-nervous system peptide  
 CC transmitters, defensins, growth factors, vasoactive peptides and  
 CC receptors, binding proteins and malignancy associated proteins. The  
 CC antisense oligonucleotides may be used in this way to treat disorders  
 CC including respiratory obstruction (especially pulmonary obstruction  
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or  
 CC surfactant hypoproduction which are associated with a disease or  
 CC condition selected from pulmonary vasoconstriction, inflammation,  
 CC allergies, asthma, impeded respiration, respiratory distress syndrome  
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary  
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),  
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,  
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide  
 CC fragments and antisense oligonucleotides used in the exemplification of  
 CC the present invention  
 CC XX  
 SQ Sequence 1742 BP; 405 A; 492 C; 466 G; 379 T; 0 U; 0 Other;  
 Query Match 43.4%; Score 624.4; DB 3; Length 1742;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAAAGCACTGCTCTGTGGTGTGCAACTGAGCGCTCTCCCA 60  
 DB 76 ATGAACCGGGAGTCCCTTTTAAAGCACTGCTCTGTGGTGTGCAACTGAGCGCTCTCCCA 135  
 QY 61 GCAGCCACTCAGGGAGAAAGTGTGTCTGTGGCAAAAAAGGGATACAGTGAACCTGACC 120  
 DB 136 GCAGCCACTCAGGGAGAAAGTGTGTCTGTGGCAAAAAAGGGATACAGTGAACCTGACC 195  
 QY 121 TGTACACTTCCCAAGAAAGATACATTCATCTGGAAGAACTCCACCAATTAAG 180  
 DB 196 TGTACACTTCCCAAGAAAGATACATTCATCTGGAAGAACTCCACCAATTAAG 255  
 QY 181 ATTCTGGAAATCAGGGCTCTCTTCTTAACCTAAGATCAATCAAGTGAATGAGCGCT 240  
 DB 256 ATTCTGGAAATCAGGGCTCTCTTCTTAACCTAAGATCAATCAAGTGAATGAGCGCT 315  
 QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAAATCTTAAG 300  
 DB 316 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAAATCTTAAG 375  
 QY 301 ATGAAGACTCAGTACTTACTTACTGTGAAGTGAAGCAAGAAAGAGAGTGCATTAAG 360  
 DB 376 ATGAAGACTCAGTACTTACTTACTGTGAAGTGAAGCAAGAAAGAGAGTGCATTAAG 435  
 QY 361 CTAGTGTTCGATTTGACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 420  
 DB 436 CTAGTGTTCGATTTGACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 495  
 QY 421 CTGACCTTGGAGAGACCCCTCTGTAGTGAAGCCCTCACTGCACTGCACTGCACTGCACTG 480  
 DB 496 CTGACCTTGGAGAGACCCCTCTGTAGTGAAGCCCTCACTGCACTGCACTGCACTGCACTG 555  
 QY 481 AAAAACAATCAGGGGGGAGAAAGCCCTCTCTGTCTCAGCTGAGCTCCAGATATGAGGC 540  
 DB 556 AAAAACAATCAGGGGGGAGAAAGCCCTCTCTGTCTCAGCTGAGCTCCAGATATGAGGC 615  
 QY 541 ACCTGCAATGCACTGTCTTGTGAGAACCAAGAAAGGTGAGTCAAAATATGACATGCTG 600  
 DB 616 ACCTGCAATGCACTGTCTTGTGAGAACCAAGAAAGGTGAGTCAAAATATGACATGCTG 675  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCCG 626  
 DB 676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701  
 RESULT 13  
 ID AA165462 standard; DNA; 1742 BP.  
 AC AA165462;  
 XX 10-DEC-2001 (first entry)  
 DT Nucleotide sequence of a human polynucleotide.  
 XX Human; receptor; DC-SIGN; dendritic cell; T lymphocyte; HIV; gp120;  
 KM C-type lectin; ICAM3; HIV entry; T cell; macrophage; HIV infection; ss.  
 OS Homo sapiens.  
 PN MO300164752-A2.  
 PD 07-SEP-2001.  
 XX 28-FEB-2001; 2001WO-US006322.  
 PF 02-MAR-2000; 2000US-00517605.  
 XX (UTNY) UNIV NEW YORK STATE.  
 PA (UTNY-) UNIV NIMEGEN.  
 PI Litman DR, Kwon D, Van Kooyk Y, Geijtenbeek T;

XX WPI; 2001-602565/68.  
 DR An antibody for the treatment or prevention of HIV-infection comprises a  
 XX gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding of  
 PT DC-SIGN due to concomitant conformational change.  
 PS Disclosure; Page 121-122; 131pp; English.  
 CC The specification describes an antibody which is specific for an  
 CC antigenic fragment of gp120. This antigenic fragment binds to DC-SIGN or  
 CC is exposed upon gp120 binding of DC-SIGN due to concomitant  
 CC conformational change. DC-SIGN is a receptor that is specifically  
 CC expressed on dendritic cells and facilitates infection of T lymphocytes  
 CC with HIV. DC-SIGN is identical to a HIV-1 gp120-binding C-type lectin. DC  
 CC -SIGN binds ICAM-3 (which is expressed constitutively on T lymphocytes)  
 CC with high affinity. The antibody of the invention inhibits the trans  
 CC enhancement of HIV entry into a T cell or macrophage facilitated by  
 CC dendritic cells. The antibody is useful to treat or prevent HIV  
 CC infection. The present sequence represents a human polynucleotide, which  
 CC is used in the course of the invention  
 XX  
 SQ Sequence 1742 BP; 405 A; 491 C; 466 G; 380 T; 0 U; 0 Other:  
 Query Match 43.4%; Score 624.4; DB 4; Length 1742;  
 Best Local Similarity 99.8%; Pred. No. 1,8e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGAACCGGGAGTCCCTTTTAAAGCACTGCTCTGTGGTGTGCAACTGAGCGCTCTCCCA 60  
 DB 76 ATGAACCGGGAGTCCCTTTTAAAGCACTGCTCTGTGGTGTGCAACTGAGCGCTCTCCCA 135  
 QY 61 GCAGCCACTCAGGGAGAAAGTGTGTCTGTGGCAAAAAAGGGATACAGTGAACCTGACC 120  
 DB 136 GCAGCCACTCAGGGAGAAAGTGTGTCTGTGGCAAAAAAGGGATACAGTGAACCTGACC 195  
 QY 121 TGTACACTTCCCAAGAAAGATACATTCATCTGGAAGAACTCCACCAATTAAG 180  
 DB 196 TGTACACTTCCCAAGAAAGATACATTCATCTGGAAGAACTCCACCAATTAAG 255  
 QY 181 ATTCTGGAAATCAGGGCTCTCTTCTTAACCTAAGATCAATCAAGTGAATGAGCGCT 240  
 DB 256 ATTCTGGAAATCAGGGCTCTCTTCTTAACCTAAGATCAATCAAGTGAATGAGCGCT 315  
 QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAAATCTTAAG 300  
 DB 316 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAAATCTTAAG 375  
 QY 301 ATGAAGACTCAGTACTTACTTACTGTGAAGTGAAGCAAGAAAGAGAGTGCATTAAG 360  
 DB 376 ATGAAGACTCAGTACTTACTTACTGTGAAGTGAAGCAAGAAAGAGAGTGCATTAAG 435  
 QY 361 CTAGTGTTCGATTTGACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 420  
 DB 436 CTAGTGTTCGATTTGACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 495  
 QY 421 CTGACCTTGGAGAGACCCCTCTGTAGTGAAGCCCTCACTGCACTGCACTGCACTGCACTG 480  
 DB 496 CTGACCTTGGAGAGACCCCTCTGTAGTGAAGCCCTCACTGCACTGCACTGCACTGCACTG 555  
 QY 481 AAAAACAATCAGGGGGGAGAAAGCCCTCTCTGTCTCAGCTGAGCTCCAGATATGAGGC 540  
 DB 556 AAAAACAATCAGGGGGGAGAAAGCCCTCTCTGTCTCAGCTGAGCTCCAGATATGAGGC 615  
 QY 541 ACCTGCAATGCACTGTCTTGTGAGAACCAAGAAAGGTGAGTCAAAATATGACATGCTG 600  
 DB 616 ACCTGCAATGCACTGTCTTGTGAGAACCAAGAAAGGTGAGTCAAAATATGACATGCTG 675  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCCG 626  
 DB 676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701

RESULT 14  
AB297021 standard; DNA; 1742 BP.  
XX  
XX AB297021:  
XX  
XX 17-OCT-2003 (first entry)  
XX  
XX Human nucleic acid sequence.  
XX  
XX Human; antisense; lung dysfunction; nasal airway dysfunction;  
XX antiinflammatory steroid; ubiquinone; antiinflammatory; antiallergic;  
XX antiallergic; hypotensive; immunosuppressive; cytoskeletal; gene therapy;  
XX antisense gene therapy; respiratory; lung; adenosine sensitivity;  
XX adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;  
XX lung inflammation; respiratory disease; ds.  
XX  
XX Homo sapiens.  
XX  
XX W0200285308-A2.  
XX  
XX 31-OCT-2002.  
XX  
XX 23-APR-2002; 2002MO-US013135.  
XX  
XX 24-APR-2001; 2001US-0286137P.  
XX  
XX (EPIG-) EPIGENESIS PHARM INC.  
XX  
XX Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;  
XX Miller S, Tang L, Shahbuddin S;  
XX WPI; 2003-229219/22.  
XX  
XX Pharmaceutical composition for treating ailments associated with impaired  
XX respiration, has oligo(s) antisense to specific gene(s) or its  
XX corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or  
XX ubiquinone.  
XX  
XX Disclousure; SEQ ID NO 1263; 872pp; English.  
XX  
XX The invention relates to a novel pharmaceutical composition, which has a  
XX first active agent comprising an oligonucleotide antisense to the  
XX initiation codon, coding region, 5' or 3' end genomic flanking regions,  
XX 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of  
XX junctions of genes encoding a polypeptide associated with lung and/or  
XX nasal airway dysfunction and a second active agent comprising an  
XX antiinflammatory steroid and ubiquinone. A composition of the invention  
XX has antiinflammatory, antiallergic, antiasthmatic, hypotensive,  
XX immunosuppressive, and cytoskeletal activity. The composition may have a  
XX use in antisense gene therapy. The composition is useful for treating or  
XX preventing a respiratory, lung or malignant disease or condition, also  
XX for enhancing the prophylactic or therapeutic respiratory effect of an  
XX antiinflammatory steroid in a subject, for reducing or depleting levels  
XX of, or reducing sensitivity to adenosine, reducing levels of adenosine  
XX receptor, producing bronchodilation, increasing levels of ubiquinone or  
XX lung surfactant in a subject's tissue, or treating bronchoconstriction,  
XX lung inflammation, lung allergies, or a respiratory disease or condition.  
XX Note: The sequence data for this patent is not represented in the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 1742 BP; 405 A; 492 C; 466 G; 379 T; 0 U; 0 Other;  
XX  
XX Query Match 43.4%; Score 624.4; DB 10; Length 1742;  
XX Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
XX Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 136 GCAGCCACTCAGGAAAGAAAGTGCTGCGGCAAAAAAGGGGATACATGTGAACCTGAC 195  
Qy 121 TGTACAGCTTCCAGAAAGAGACATACAAATTCATGTGAAAACTCAACCATGATTAAG 180  
Db 196 TGTACAGCTTCCAGAAAGAGACATACAAATTCATGTGAAAACTCAACCATGATTAAG 255  
Qy 181 ATTCTGGAAATCAGGGCTCTTCTTAATAAGTTCATCAAGCTGAATGATCGGCT 240  
Db 256 ATTCTGGAAATCAGGGCTCTTCTTAATAAGTTCATCAAGCTGAATGATCGGCT 315  
Qy 241 GATTCAGAAAGAGCTTTGGGACCAAGAACTTCCCTGTATCATCAAGATTTAG 300  
Db 316 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGTATCATCAAGATTTAG 375  
Qy 301 ATAGAGACTCAGATATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 360  
Db 376 ATAGAGACTCAGATATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 435  
Qy 361 CTAGTGTGGATTGATCTGCCAATCTGTACACCCACTGCTTCAGGGGACAGCTGACC 420  
Db 436 CTAGTGTGGATTGATCTGCCAATCTGTACACCCACTGCTTCAGGGGACAGCTGACC 495  
Qy 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATATGATGAGTCCAGGGGT 480  
Db 496 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATATGATGAGTCCAGGGGT 555  
Qy 481 AAAACATACAGGGGGGAAAGACCTCTCGTGTCTCAGCTGAGAGCTCCAGATATGAGC 540  
Db 556 AAAACATACAGGGGGGAAAGACCTCTCGTGTCTCAGCTGAGAGCTCCAGATATGAGC 615  
Qy 541 ACCTGACATGCACTGTCTTGAGAAACAGAAAGTGGAGTTCAAAATAGACATCGTG 600  
Db 616 ACCTGACATGCACTGTCTTGAGAAACAGAAAGTGGAGTTCAAAATAGACATCGTG 675  
Qy 601 GTGCTAGCTTTCAGAAAGGCTCCCG 626  
Db 676 GTGCTAGCTTTCAGAAAGGCTCCAG 701

RESULT 15  
AD131687  
ID AD131687 standard; cDNA; 1742 BP.  
XX  
XX AD131687:  
XX  
XX 17-JUN-2004 (first entry)  
XX  
XX Human cDNA #1013.  
XX  
XX Human; gene; ss; immunological response; immunopathological condition;  
XX Crohn's disease; asthma; ulcerative colitis; hypersplenophilia;  
XX irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;  
XX acute monocytic leukemia; antiinflammatory; antiasthmatic; antulcer;  
XX osteopathic; antirheumatic; antirheumatic; cytoskeletal.  
XX  
XX Homo sapiens.  
XX  
XX US6607879-B1.  
XX  
XX 19-AUG-2003.  
XX  
XX 09-FEB-1998; 98US-00023655.  
XX  
XX 09-FEB-1998; 98US-00023655.  
XX  
XX (INCY-) INCYTE CORP.  
XX  
XX Cocks BG, Stuart SG, Seilhamer JI;  
XX WPI; 2003-895307/82.  
XX  
XX A composition comprising a plurality of cDNAs, useful for detecting

PT altered expression of genes in an immunological response or for  
PT diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma  
PT or osteoarthritis.

PS Claim 1, SEQ ID NO 1013, 50bp, English.

XX The invention relates to a composition comprising a plurality of cDNAs  
XX for detecting the altered expression of genes in an immunological  
CC response. The invention also relates to a method of diagnosing or  
CC monitoring the treatment of an immunopathological condition in a sample,  
CC comprising obtaining nucleic acids from a sample, contacting the nucleic  
CC acids of the sample with an array comprising the plurality of cDNAs under  
CC conditions to form one or more hybridisation complexes, detecting the  
CC hybridisation complexes and comparing the levels of the detected  
CC hybridisation complexes with the level of hybridisation complexes  
CC detected in a non-diseased sample, where an altered level of the detected  
CC hybridisation complexes correlates with the presence of an  
CC immunopathological condition. Also disclosed are an expression profile  
CC comprising a microarray and a plurality of detectable complexes and a  
CC method for identifying a plurality of polynucleotide probes. The cDNAs  
CC are useful as hybridisable array elements in a microarray for monitoring  
CC the expression of target polynucleotides. The microarray can be used in  
CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,  
CC ulcerative colitis, hypersplenophilia, irritable bowel syndrome,  
CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in  
CC identifying agents for the treatment of the diseases. The microarray may  
CC also be used in drug discovery and development, toxicological and  
CC carcinogenicity studies, forensics or pharmacogenomics. The composition  
CC may also be used in purification of a subpopulation of mRNAs, cDNAs or  
CC genomic fragments. This sequence represents a human cDNA of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification but was obtained in electronic format directly  
CC from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 1742 BP; 405 A; 491 C; 466 G; 380 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 11; Length 1742;

Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGGAGTCCCTTTAGGCACTTCTGTGTGCAACTGGCGCTCTCCCA 60  
Db 76 ATGAACCGGGGAGTCCCTTTAGGCACTTCTGTGTGCAACTGGCGCTCTCCCA 135  
QY 61 GCAGCCCACTCAGGGAAGAAAGTGTGTGGCAAAAAAGGGAATACAGTGAAC 120  
Db 136 GCAGCCCACTCAGGGAAGAAAGTGTGTGGCAAAAAAGGGAATACAGTGAAC 195  
QY 121 TGTACAGCTTCCAGAGAAGACATACATTCACCTGGAATACTCCACCAAGATAAG 180  
Db 196 TGTACAGCTTCCAGAGAAGACATACATTCACCTGGAATACTCCACCAAGATAAG 255  
QY 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCCAGCTGAATGCGCT 240  
Db 256 ATTCTGGGAATCAGGGCTCTCTTAAAGTCCATCCAGCTGAATGCGCT 315  
QY 241 GACTCAAGAGAAGCCTTTGGAGCAAGAAATCTCCCTGATCATTAAGATTTAAG 300  
Db 316 GACTCAAGAGAAGCCTTTGGAGCAAGAAATCTCCCTGATCATTAAGATTTAAG 375  
QY 301 ATAGAAGACAGATTAATCATCTGTAGTGAAGAGACCAAGAGAAGAGTGAATTTG 360  
Db 376 ATAGAAGACAGATTAATCATCTGTAGTGAAGAGACCAAGAGAAGAGTGAATTTG 435  
QY 361 CTAGTGTTCGATTTGACTGCAACTGCAACCACTGCTTCAAGGGCAGAGCTGACC 420  
Db 436 CTAGTGTTCGATTTGACTGCAACTGCAACCACTGCTTCAAGGGCAGAGCTGACC 495  
QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAAGTGAATGAGATCAAGGGGT 480  
Db 496 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAAGTGAATGAGATCAAGGGGT 555  
QY 481 AAAAACAATACAGGGGGGAAGACCTCTCCGTGTCTAGCTGAGCTCAAGATATGTGC 540

Db 556 AAAAACAATACAGGGGGGAAGACCTCTCCGTGTCTAGCTGAGCTCAAGATATGTGC 615  
QY 541 ACCTGACATGCACTGTCTTTCAGAGAACCAAGAGAAGTGGAGTTCAAAATAGACATGCTG 600  
Db 616 ACCTGACATGCACTGTCTTTCAGAGAACCAAGAGAAGTGGAGTTCAAAATAGACATGCTG 675  
QY 601 GTGCTAGCTTTCAGAGAAGGCTCCGG 626  
Db 676 GTGCTAGCTTTCAGAGAAGGCTCCAG 701

Search completed: September 20, 2005, 10:37:09  
Job time : 870 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 20, 2005, 08:29:41 ; Search time 298 Seconds  
(without alignments)  
7906.841 Million cell updates/sec

Title: US-09-936-702-4  
Perfect score: 1440  
Sequence: 1 atgaaccgggagctccctt.....gtccagcagcacaataactcg 1440

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapect 1.0

Searched: 1202784 seqs, 818138359 residues  
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: Issued\_Patents\_NA:\*  
2: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	625	43.4	1213	6 5223418-1	Patent No. 5223418
2	625	43.4	1213	6 5223418-1	Patent No. 5223418
3	624.4	43.4	1415	5 PCT-US92-08090-4	Sequence 4, Appl1
4	624.4	43.4	1421	5 PCT-US92-08090-3	Sequence 3, Appl1
5	624.4	43.4	1448	5 PCT-US92-08090-2	Sequence 2, Appl1
6	624.4	43.4	1742	3 US-09-517-605-7	Sequence 7, Appl1
7	624.4	43.4	1742	4 US-09-023-655-1013	Sequence 1013, Ap
8	624.4	43.4	2465	5 PCT-US92-08090-1	Sequence 1, Appl1
9	622.8	43.2	1304	2 US-08-284-391B-28	Sequence 28, Appl1
10	622.8	43.2	1304	3 US-09-218-950-1	Sequence 28, Appl1
11	622.8	43.2	1304	4 US-08-394-388A-28	Sequence 28, Appl1
12	622.8	43.2	1389	2 US-08-417-495-2	Sequence 2, Appl1
13	622.8	43.2	1389	2 US-08-284-391B-2	Sequence 2, Appl1
14	622.8	43.2	1389	3 US-09-218-950-2	Sequence 2, Appl1
15	622.8	43.2	1389	4 US-08-394-388A-2	Sequence 2, Appl1
16	622.8	43.2	1389	5 PCT-US92-01785-2	Sequence 2, Appl1
17	622.8	43.2	1389	5 PCT-US95-00454-2	Sequence 2, Appl1
18	622.8	43.2	1599	2 US-08-417-495-3	Sequence 3, Appl1
19	622.8	43.2	1599	2 US-08-284-391B-3	Sequence 3, Appl1
20	622.8	43.2	1599	3 US-09-218-950-3	Sequence 3, Appl1
21	622.8	43.2	1599	4 US-08-394-388A-3	Sequence 3, Appl1
22	622.8	43.2	1599	5 PCT-US92-01785-3	Sequence 3, Appl1
23	622.8	43.2	1599	5 PCT-US95-00454-3	Sequence 3, Appl1
24	622.8	43.2	1728	2 US-08-417-495-1	Sequence 1, Appl1
25	622.8	43.2	1728	2 US-08-284-391B-1	Sequence 1, Appl1
26	622.8	43.2	1728	3 US-09-218-950-1	Sequence 1, Appl1
27	622.8	43.2	1728	4 US-08-394-388A-1	Sequence 1, Appl1

28	622.8	43.2	1728	5 PCT-US92-01785-1	Sequence 1, Appl1
29	622.8	43.2	1728	5 PCT-US95-00454-1	Sequence 1, Appl1
30	622.8	43.2	1742	3 US-08-466-368-3	Sequence 3, Appl1
31	622.8	43.2	1742	4 US-08-470-998-1	Sequence 1, Appl1
32	622.8	43.2	1742	4 US-08-328-500-8	Sequence 8, Appl1
33	622.8	43.2	2589	4 US-08-472-888A-5	Sequence 5, Appl1
34	621.2	43.1	1273	3 US-08-466-368-1	Sequence 1, Appl1
35	621.2	43.1	1273	4 US-08-328-500-1	Sequence 1, Appl1
36	621.2	43.1	1416	1 US-08-236-311-2	Sequence 2, Appl1
37	621.2	43.1	1416	3 US-08-457-918-2	Sequence 2, Appl1
38	621.2	43.1	1416	4 US-10-157-408-2	Sequence 2, Appl1
39	621.2	43.1	2940	2 US-08-418-848A-3	Sequence 3, Appl1
40	621.2	43.1	3426	2 US-08-418-848A-2	Sequence 2, Appl1
41	621.2	43.1	3480	2 US-08-418-848A-4	Sequence 4, Appl1
42	621.2	43.1	3721	2 US-08-418-848A-5	Sequence 5, Appl1
43	621.2	43.1	3993	2 US-08-418-848A-6	Sequence 6, Appl1
44	621.2	43.1	4059	2 US-08-418-848A-7	Sequence 7, Appl1
45	621.2	43.1	4632	2 US-08-418-848A-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1  
5223418-1  
; Patent No. 5223418  
; APPLICANT: ARCURI, EDWARD J.;BRAUNER, MARY E.; DONOVAN, MARY  
; J.; GERBER, ROBERT G.; KELLER, JOHN A.  
; TITLE OF INVENTION: METHOD OF IMPROVING THE YIELD OF  
; HETEROLOGOUS PROTEINS PRODUCED BY STREPTOMYCES LIVIDANS  
; NUMBER OF SEQUENCES: 2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/589,979  
; FILING DATE: 28-SEP-1990  
; SEQ ID NO:1:  
; LENGTH: 1213  
5223418-1

Query Match	43.4%	Score 625;	DB 6;	Length 1213;
Best Local Similarity	97.7%	Pred. No. 3.7e-138;		
Matches 634;	Conservative 0;	Mismatches 15;	Indels 0;	Gaps 0;
QY	1	ATGAACGGGAGATCCCTTTTAAAGCACTTGCTGTGCTGCACTGGCGCTCCCA	60	
DB	79	ATGAACGGGAGATCCCTTTTAAAGCACTTGCTGTGCTGCACTGGCGCTCCCA	138	
QY	61	GCAGCCACTCAGGAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAATGACC	120	
DB	139	GCAGCCACTCAGGAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAATGACC	198	
QY	121	TGTACAGCTTCCAGAAAGCATCAATTCATGGAATAAATCCCAACGATTAAG	180	
DB	199	TGTACAGCTTCCAGAAAGCATCAATTCATGGAATAAATCCCAACGATTAAG	258	
QY	181	ATTCTGGGAATCAGGGCTCTCTTAAAGTCAATCCAAAGTGAATGCGCT	240	
DB	259	ATTCTGGGAATCAGGGCTCTCTTAAAGTCAATCCAAAGTGAATGCGCT	318	
QY	241	GACTCAAGAAAGCTTTGGGACCAAGAAATCCCTCCCTGATCATCAAGATTTAAG	300	
DB	319	GACTCAAGAAAGCTTTGGGACCAAGAAATCCCTCCCTGATCATCAAGATTTAAG	378	
QY	301	ATGAAGACCTCAGATCTTAACATCTGTGAAGTGGAGACCAAGAGAGAGGCAATTG	360	
DB	379	ATGAAGACCTCAGATCTTAACATCTGTGAAGTGGAGACCAAGAGAGAGGCAATTG	438	
QY	361	CTAGTGTGGATGATGCTGCACTGCAACCACTGCTTCAAGGGAGAGCTGACC	420	
DB	439	CTAGTGTGGATGATGCTGCACTGCAACCACTGCTTCAAGGGAGAGCTGACC	498	
QY	421	CTGACCTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATGTATGAGTCCAAAGGGT	480	
DB	499	CTGACCTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATGTATGAGTCCAAAGGGT	558	

```

QY      481  AAAAATACAGAGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGGC 540
        |||
Db      559  AAAAATACAGAGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGGC 618

QY      541  ACCTGACATGCACTGTCTTGCAGAACCAAGAAAGGTGAGATTCAAAATAGACATGTG 600
        |||
Db      619  ACCTGACATGCACTGTCTTGCAGAACCAAGAAAGGTGAGATTCAAAATAGACATGTG 678

QY      601  GTGCTAGCTTCCAGAGGCTCCGAGAGTGCGCGGTAGTGGGGAGCG 649
        |||
Db      679  GTGCTAGCTTCCAGAGGCTCCGAGATAGTCTATAGAAAGAGGCG 727

```

```

RESULT 2
5223418-1
; Patent No. 5223418
; APPLICANT: ARCURI, EDWARD J.; BRANNER, MARY E.; DONOVAN, MARY
; J.; GERBER, ROBERT G.; KELLER, JOHN A.
; TITLE OF INVENTION: METHOD OF IMPROVING THE YIELD OF
; HETEROLOGOUS PROTEINS PRODUCED BY STREPTOMYCES LIVIDANS
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/589,979
; FILING DATE: 28-SEP-1990
; SEQ ID NO:1:
; LENGTH: 1213
5223418-1

```

```

Query Match      43.4%; Score 625; DB 6; Length 1213;
Best Local Similarity 97.7%; Pred. No. 3,7e-138;
Matches 634; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      1  ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGTGCTGCACTGGCGCTCTCCCA 60
        |||
Db      79  ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGTGCTGCACTGGCGCTCTCCCA 138

QY      61  GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120
        |||
Db      139  GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 198

QY      121  TGTACAGCTTCCAGAGAAAGAGATCAATTCACCTGGAAAACTCCACGATTAAG 180
        |||
Db      199  TGTACAGCTTCCAGAGAAAGAGATCAATTCACCTGGAAAACTCCACGATTAAG 258

QY      181  ATTCTGGGAATCAGGGCTCTTCTTAATAAGTCCATCAAGCTGAATGATCGGCT 240
        |||
Db      259  ATTCTGGGAATCAGGGCTCTTCTTAATAAGTCCATCAAGCTGAATGATCGGCT 318

QY      241  GACTCAAGAAAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 300
        |||
Db      319  GACTCAAGAAAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 378

QY      301  ATGAAGACTCAATATTAATCTGTGAAGTGAAGACCAAGAAAGAGGTGCAATTG 360
        |||
Db      379  ATGAAGACTCAATATTAATCTGTGAAGTGAAGACCAAGAAAGAGGTGCAATTG 438

QY      361  CTAGTGTTCGATTGACTGCAACTGTGACACCACTGCTTCAAGGGGAGAGCTGACC 420
        |||
Db      439  CTAGTGTTCGATTGACTGCAACTGTGACACCACTGCTTCAAGGGGAGAGCTGACC 498

QY      421  CTGACCTTGAAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGT 480
        |||
Db      499  CTGACCTTGAAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGT 558

QY      481  AAAAATACAGAGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGGC 540
        |||
Db      559  AAAAATACAGAGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGGC 618

QY      541  ACCTGACATGCACTGTCTTGCAGAACCAAGAAAGGTGAGATTCAAAATAGACATGTG 600
        |||
Db      619  ACCTGACATGCACTGTCTTGCAGAACCAAGAAAGGTGAGATTCAAAATAGACATGTG 678

```

```

QY      601  GTGCTAGCTTCCAGAGGCTCCGAGAGTGCGCGGTAGTGGGGAGCG 649
        |||
Db      679  GTGCTAGCTTCCAGAGGCTCCGACATAGTCTATAGAAAGAGGCG 727

```

```

RESULT 3
PCT-US92-08090-4
; Sequence 4, Application PC/TUS9208090
; GENERAL INFORMATION:
; APPLICANT: Tang, J. N.
; TITLE OF INVENTION: Fusion Protein Genes for Treatment of
; TITLE OF INVENTION: AIDS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kipatrick & Cody
; STREET: 100 Peachtree Street
; City: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08090
; FILING DATE: 19920922
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patricia U.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMR#129
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1415 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: epithelial
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..6
; OTHER INFORMATION: /note= "Restriction site"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1275..1280
; OTHER INFORMATION: /note= "Restriction site"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1410..1415
; OTHER INFORMATION: /note= "Restriction site"
PCT-US92-08090-4

```

```

Query Match      43.4%; Score 624.4; DB 5; Length 1415;
Best Local Similarity 99.8%; Pred. No. 5.4e-138;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGTGCTGCAACTGGCGCTCTCCCA 60
        |||
Db      81  ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGTGCTGCAACTGGCGCTCTCCCA 140

QY      61  GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120
        |||
Db      141  GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 200

```

QY 121 TGTACAGCTTCCGAGAGAGATCAATTCATCTGAAAGAACTCCAAACAGATTAAG 180  
 Db 201 TGTACAGCTTCCGAGAGAGATCAATTCATCTGAAAGAACTCCAAACAGATTAAG 260  
 QY 181 ATTCTGGGAAATCAGGGCTCTCTTCTTAACCTAAAGGTCATCCAGAGTGAATGATCGGCT 240  
 Db 261 ATTCTGGGAAATCAGGGCTCTCTTCTTAACCTAAAGGTCATCCAGAGTGAATGATCGGCT 320  
 QY 241 GACTCAAGAAAGAGCTTTTGGAGCCAGGAACTTCCCTGATCATCAAGAACTTAAG 300  
 Db 321 GACTCAAGAAAGAGCTTTTGGAGCCAGGAACTTCCCTGATCATCAAGAACTTAAG 380  
 QY 301 ATAGAAAGCTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGAGTGAATG 360  
 Db 381 ATAGAAAGCTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGAGTGAATG 440  
 QY 361 CTAGGTGTTGGATTGACTGCTCAACTGACACCCCACTGCTTCAGGGGCAAGGCTGACC 420  
 Db 441 CTAGGTGTTGGATTGACTGCTCAACTGACACCCCACTGCTTCAGGGGCAAGGCTGACC 500  
 QY 421 CTAGCTTGGAGAGCCCTCTGTGATGAGCCCTCAGTGCAGAACTGAGATCCAAAGGGGT 480  
 Db 501 CTAGCTTGGAGAGCCCTCTGTGATGAGCCCTCAGTGCAGAACTGAGATCCAAAGGGGT 560  
 QY 481 AAAAACAATACAGGGGGGAGAAACCTCTCCGTCTCAGCTGAGCTCCAGATATGAGC 540  
 Db 561 AAAAACAATACAGGGGGGAGAAACCTCTCCGTCTCAGCTGAGCTCCAGATATGAGC 620  
 QY 541 ACCTGACATGACATGCTCTTGCAGAACCAAGAAAGGTGAGATTCAAAATATGACATGCTG 600  
 Db 621 ACCTGACATGACATGCTCTTGCAGAACCAAGAAAGGTGAGATTCAAAATATGACATGCTG 680  
 QY 601 GTGCTAGCTTCCAGAGGCTCCGG 626  
 Db 681 GTGCTAGCTTCCAGAGGCTCCAG 706

## RESULT 4

PCT-US92-08090-3  
 ; Sequence 3, Application PC/TUS9208090  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, J. N.  
 ; TITLE OF INVENTION: Fusion Protein Genes for Treatment of  
 ; TITLE OF INVENTION: Aids  
 ; NUMBER OF SEQUENCES: 15  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Kilpatrick & Cody  
 ; STREET: 100 Peachtree Street  
 ; CITY: Atlanta  
 ; STATE: Georgia  
 ; COUNTRY: U.S.  
 ; ZIP: 30303  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US92/08090  
 ; FILING DATE: 19920922  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Pabst, Patricia L.  
 ; REGISTRATION NUMBER: 31,284  
 ; REFERENCE/DOCKET NUMBER: OMRP129  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 404-572-6508  
 ; TELEFAX: 404-572-6555  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1421 base pairs  
 ; TYPE: NUCLEIC ACID

; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; HYPOTHEICAL: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Homo sapien  
 ; TISSUE TYPE: Epithelial  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: 1..6  
 ; OTHER INFORMATION: /note= "Restriction site"  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: 1275..1280  
 ; OTHER INFORMATION: /note= "Restriction site"  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: 1416..1421  
 ; OTHER INFORMATION: /note= "Restriction site"  
 ; PCT-US92-08090-3

Query Match 43.4%; Score 624.4; DB 5; Length 1421;  
 Best Local Similarity 99.8%; Pred. No. 5, 4e-138;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACCGGGAGATCCCTTTTATGACATGCTTCTGTGCTGCACTGGGCTCTCCCA 60  
 Db 81 ATGACCGGGAGATCCCTTTTATGACATGCTTCTGTGCTGCACTGGGCTCTCCCA 140  
 QY 61 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAGGGGATACATGAACTGACC 120  
 Db 141 GCAGCCACTCAGGAGAAAGAGTGTGTGGGCAAAAAGGGGATACATGAACTGACC 200  
 QY 121 TGTACAGCTTCCGAGAGAGATCAATTCATCTGAAAGAACTCCAAACAGATTAAG 180  
 Db 201 TGTACAGCTTCCGAGAGAGATCAATTCATCTGAAAGAACTCCAAACAGATTAAG 260  
 QY 181 ATTCTGGGAAATCAGGGCTCTCTTCTTAACCTAAAGGTCATCCAGAGTGAATGATCGGCT 240  
 Db 261 ATTCTGGGAAATCAGGGCTCTCTTCTTAACCTAAAGGTCATCCAGAGTGAATGATCGGCT 320  
 QY 241 GACTCAAGAAAGAGCTTTTGGAGCCAGGAACTTCCCTGATCATCAAGAACTTAAG 300  
 Db 321 GACTCAAGAAAGAGCTTTTGGAGCCAGGAACTTCCCTGATCATCAAGAACTTAAG 380  
 QY 301 ATAGAAAGCTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGAATG 360  
 Db 381 ATAGAAAGCTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGAATG 440  
 QY 361 CTAGGTGTTGGATTGACTGCTCAACTGACACCCCACTGCTTCAGGGGCAAGGCTGACC 420  
 Db 441 CTAGGTGTTGGATTGACTGCTCAACTGACACCCCACTGCTTCAGGGGCAAGGCTGACC 500  
 QY 421 CTAGCTTGGAGAGCCCTCTGTGATGAGCCCTCAGTGCAGAACTGAGATCCAAAGGGGT 480  
 Db 501 CTAGCTTGGAGAGCCCTCTGTGATGAGCCCTCAGTGCAGAACTGAGATCCAAAGGGGT 560  
 QY 481 AAAAACAATACAGGGGGGAGAAACCTCTCCGTCTCAGCTGAGCTCCAGATATGAGC 540  
 Db 561 AAAAACAATACAGGGGGGAGAAACCTCTCCGTCTCAGCTGAGCTCCAGATATGAGC 620  
 QY 541 ACCTGACATGACATGCTCTTGCAGAACCAAGAAAGGTGAGATTCAAAATATGACATGCTG 600  
 Db 621 ACCTGACATGACATGCTCTTGCAGAACCAAGAAAGGTGAGATTCAAAATATGACATGCTG 680  
 QY 601 GTGCTAGCTTCCAGAGGCTCCGG 626  
 Db 681 GTGCTAGCTTCCAGAGGCTCCAG 706

## RESULT 5

PCT-US92-08090-2

```

; Sequence 2, Application PC/TUS9208090
; GENERAL INFORMATION:
; APPLICANT: Tang, J. N.
; TITLE OF INVENTION: Fusion Protein Genes for Treatment of
; TITLE OF INVENTION: Aids
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08090
; FILING DATE: 19920922
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF129
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1448 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Epithelial
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..6
; OTHER INFORMATION: /note= "Restriction site"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1275..1280
; OTHER INFORMATION: /note= "Restriction site"
; NAME/KEY: misc_feature
; LOCATION: 1444..1448
; OTHER INFORMATION: /note= "Restriction site"
; PCT-US92-08090-2

Query Match 43.4%; Score 624.4; DB 5; Length 1448;
Best Local Similarity 99.8%; Pred. No. 5.4e-138;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGCTGCAACTGGGGCTCCCTCCCA 60
DB 81 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGCTGCAACTGGGGCTCCCTCCCA 140
QY 61 GCAGCCACTCAGGAAAGAAAGTGTGCTGGCAAAAAAGGGATACAGTGAATGACC 120
DB 141 GCAGCCACTCAGGAAAGAAAGTGTGCTGGCAAAAAAGGGATACAGTGAATGACC 200
QY 121 TGTACAGCTTCCCAAGAAAGAGCATATTCACCTGGAATACTCCAAACAGATTAAG 180
DB 201 TGTACAGCTTCCCAAGAAAGAGCATATTCACCTGGAATACTCCAAACAGATTAAG 260
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACATAAGTTCATCCAGCTGAATGATCGGCT 240

```

```

DB 261 ATTCTGGGAATCAGGGCTCTTCTTAACATAAGTTCATCCAGCTGAATGATCGGCT 320
QY 241 GACTCAAGAAAGACCTTTGGAGCAAGAAATCTCCCTGATCATCAAGATCTTAAG 300
DB 321 GACTCAAGAAAGACCTTTGGAGCAAGAAATCTCCCTGATCATCAAGATCTTAAG 380
QY 301 ATGAAGACTCAGATATCTTAATCTGTGAATGAGAGACAGAAAGAGAGGTGCAATTG 360
DB 381 ATGAAGACTCAGATATCTTAATCTGTGAATGAGAGACAGAAAGAGAGGTGCAATTG 440
QY 361 CTAGTGTGGATTGATGCTGCAACTGTGACACCACTGCTCAGGGGAGAGCTGACC 420
DB 441 CTAGTGTGGATTGATGCTGCAACTGTGACACCACTGCTCAGGGGAGAGCTGACC 500
QY 421 CTGACCTTGAAGACCCCTGATGATGACCCCTCAGTGAATGAGAGTCAAGAGGT 480
DB 501 CTGACCTTGAAGACCCCTGATGATGACCCCTCAGTGAATGAGAGTCAAGAGGT 560
QY 481 AAAACATACAGGGGGGAAAGACCTCTCGTGTCTCAGCTGAGCTCCAGATATGTGC 540
DB 561 AAAACATACAGGGGGGAAAGACCTCTCGTGTCTCAGCTGAGCTCCAGATATGTGC 620
QY 541 ACCTGACATGCACTGTCTTGCAAGACAGAAAGGTGAGTTCAAAATTAGACATCGTG 600
DB 621 ACCTGACATGCACTGTCTTGCAAGACAGAAAGGTGAGTTCAAAATTAGACATCGTG 680
QY 601 GTGCTAGCTTTCAGAAAGCCTCCG 626
DB 681 GTGCTAGCTTTCAGAAAGCCTCCG 706

RESULT 6
US-09-517-605-7
; Sequence 7, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Liltman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geljtenbeck, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517, 605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1742
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-517-605-7

Query Match 43.4%; Score 624.4; DB 3; Length 1742;
Best Local Similarity 99.8%; Pred. No. 5.7e-138;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGCTGCAACTGGGGCTCCCTCCCA 60
DB 76 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGCTGCAACTGGGGCTCCCTCCCA 135
QY 61 GCAGCCACTCAGGAAAGAAAGTGTGCTGGCAAAAAAGGGATACAGTGAATGACC 120
DB 136 GCAGCCACTCAGGAAAGAAAGTGTGCTGGCAAAAAAGGGATACAGTGAATGACC 195
QY 121 TGTACAGCTTCCCAAGAAAGAGCATATTCACCTGGAATACTCCAAACAGATTAAG 180
DB 196 TGTACAGCTTCCCAAGAAAGAGCATATTCACCTGGAATACTCCAAACAGATTAAG 255
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACATAAGTTCATCCAGCTGAATGATCGGCT 240
DB 256 ATTCTGGGAATCAGGGCTCTTCTTAACATAAGTTCATCCAGCTGAATGATCGGCT 315

```



SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/08090  
FILING DATE: 19920922  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRF129  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-572-6508  
TELEFAX: 404-572-6555  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2465 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapien  
TISSUE TYPE: Epithelial  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: 1..6  
OTHER INFORMATION: /note= "Restriction site"  
NAME/KEY: misc.feature  
LOCATION: 2460..2465  
OTHER INFORMATION: /note= "Restriction site"  
PCT-US92-08090-1

Query Match 43.4%; Score 624.4; DB 5; Length 2465;  
Best Local Similarity 99.8%; Pred. No. 6.3e-138;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGGAGTCCCTTTTAGGCACTTGCTGCTGCAACTGGCGCTCTCCCA 60  
DB 81 ATGAACCGGGGAGTCCCTTTTAGGCACTTGCTGCTGCTGCAACTGGCGCTCTCCCA 140  
QY 61 GCAGCCCTCAGGAGAAAGTGTGCTGGGCAAAAAGGGATTAAGTGAATGACC 120  
DB 141 GCAGCCCTCAGGAGAAAGTGTGCTGGGCAAAAAGGGATTAAGTGAATGACC 200  
QY 121 TGTACAGCTTCCGAGAAAGAGATCAATTCACCTGGAATACTCAACGATTAAG 180  
DB 201 TGTACAGCTTCCGAGAAAGAGATCAATTCACCTGGAATACTCAACGATTAAG 260  
QY 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCCAGCTGAATGATCGGCT 240  
DB 261 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCCAGCTGAATGATCGGCT 320  
QY 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 321 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 380  
QY 301 ATGAAGACTCAGATTACTTACATCTGGAAGTGAAGACCAAGAGAGAGTGCATTG 360  
DB 381 ATGAAGACTCAGATTACTTACATCTGGAAGTGAAGACCAAGAGAGAGTGCATTG 440  
QY 361 CTAGTGTGGATTTAGTGCCTGCACTGACACCCAGCTGCTTCAGGGGCAAGGCTGACC 420  
DB 441 CTAGTGTGGATTTAGTGCCTGCACTGACACCCAGCTGCTTCAGGGGCAAGGCTGACC 500  
QY 421 CTGACCTTGAAGAGCCCCCTGTGTAGTACCCCTCACTGCAATGTAGAGTCCAGGGGT 480  
DB 501 CTGACCTTGAAGAGCCCCCTGTGTAGTACCCCTCACTGCAATGTAGAGTCCAGGGGT 560  
QY 481 AAAAACAATACAGGGGGGAGAAACCTCTCCGCTGTCTAGCTGAGAGCTCCAGGATATGGC 540  
DB 561 AAAAACAATACAGGGGGGAGAAACCTCTCCGCTGTCTAGCTGAGAGCTCCAGGATATGGC 620

QY 541 ACCTGACATGCACTGTCTTGAGAAACGAAAGAGTGGAGTTCAAAATAGACATCTG 600  
DB 621 ACCTGACATGCACTGTCTTGAGAAACGAAAGAGTGGAGTTCAAAATAGACATCTG 680  
QY 601 GTGCTAGCTTCCAGAAAGGCTCCGG 626  
DB 681 GTGCTAGCTTCCAGAAAGGCTCCAG 706

## RESULT 9

US-08-284-391B-28  
Sequence 28, Application US/08284391B  
Patent No. 5851828  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391B  
FILING DATE: 02-AUG-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1304 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-284-391B-28

Query Match 43.2%; Score 622.8; DB 2; Length 1304;  
Best Local Similarity 99.7%; Pred. No. 1.3e-137;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGGAGTCCCTTTTAGGCACTTGCTGCTGCAACTGGCGCTCTCCCA 60  
DB 111 ATGAACCGGGGAGTCCCTTTTAGGCACTTGCTGCTGCTGCAACTGGCGCTCTCCCA 170  
QY 61 GCAGCCCTCAGGAGAAAGTGTGCTGGGCAAAAAGGGATTAAGTGAATGACC 120  
DB 171 GCAGCCCTCAGGAGAAAGTGTGCTGGGCAAAAAGGGATTAAGTGAATGACC 230

QY 121 TGTACAGCTTCCGAGAGAGAGATCAATTCACCTGGAAGAACTCCACAGATAAG 180  
 DB 231 TGTACAGCTTCCGAGAGAGAGATCAATTCACCTGGAAGAACTCCACAGATAAG 290  
 QY 181 ATTCTGGGAATCAGGGCTCTCTTTAACTTAAGCTTCACAGTGAATGATCGGCT 240  
 DB 291 ATTCTGGGAATCAGGGCTCTCTTTAACTTAAGCTTCACAGTGAATGATCGGCT 350  
 QY 241 GACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 DB 351 GACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 410  
 QY 301 ATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 DB 411 ATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 470  
 QY 361 CTAGCTTGGAGATGAGTCCCACTGACACCACTGCTTCAAGGGGAGAGCTGACC 420  
 DB 471 CTAGCTTGGAGATGAGTCCCACTGACACCACTGCTTCAAGGGGAGAGCTGACC 530  
 QY 421 CTGACCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 DB 531 CTGACCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 590  
 QY 481 AAAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
 DB 591 AAAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 650  
 QY 541 ACCTGAGATGACCTGCTTTCAGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
 DB 651 ACCTGAGATGACCTGCTTTCAGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 710  
 QY 601 GTGCTAGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 626  
 DB 711 GTGCTAGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 736

## RESULT 10

US-09-218-950-28

Sequence 28, Application US/09218950

Patent No. 6284240

GENERAL INFORMATION:

APPLICANT: Seed, Brian

APPLICANT: Banapour, Babak

APPLICANT: Romeo, Charles

APPLICANT: Kolanus, Waldemar

TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED

TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR - BEARING CELLS

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESSEE: Clark &amp; Elbing LLP

STREET: 176 Federal Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/218,950

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/284,391

FILING DATE: 02-AUG-1994

APPLICATION NUMBER: 08/195,395

FILING DATE: 14-FEB-1994

APPLICATION NUMBER: 07/847,566

FILING DATE: 06-MAR-1992

APPLICATION NUMBER: 07/665,961  
 FILING DATE: 07-MAR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Elbing, Karen L.  
 REGISTRATION NUMBER: 35,238  
 REFERENCE/DOCKET NUMBER: 00786/247001  
 TELEPHONE: 617-428-0200  
 TELEFAX: 617-428-7045  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 28:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1304 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-09-218-950-28

Query Match 43.2%; Score 622.8; DB 3; Length 1304;  
 Best Local Similarity 99.7%; Pred. No. 1.3e-137;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAACCGGGAGATCCCTTTTAAAGCACTTGCTTCTGCTGCACTGCGCTCTCCCA 60  
 DB 111 ATGAAACCGGGAGATCCCTTTTAAAGCACTTGCTTCTGCTGCACTGCGCTCTCCCA 170  
 QY 61 GCAGCACTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
 DB 171 GCAGCACTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 230  
 QY 121 TGTACAGCTTCCGAGAGAGAGATCAATTCACCTGGAAGAACTCCACAGATAAG 180  
 DB 231 TGTACAGCTTCCGAGAGAGAGATCAATTCACCTGGAAGAACTCCACAGATAAG 290  
 QY 181 ATTCTGGGAATCAGGGCTCTCTTTAACTTAAGCTTCACAGTGAATGATCGGCT 240  
 DB 291 ATTCTGGGAATCAGGGCTCTCTTTAACTTAAGCTTCACAGTGAATGATCGGCT 350  
 QY 241 GACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 DB 351 GACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 410  
 QY 301 ATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 DB 411 ATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 470  
 QY 361 CTAGCTTGGAGATGAGTCCCACTGACACCACTGCTTCAAGGGGAGAGCTGACC 420  
 DB 471 CTAGCTTGGAGATGAGTCCCACTGACACCACTGCTTCAAGGGGAGAGCTGACC 530  
 QY 421 CTGACCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 DB 531 CTGACCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 590  
 QY 481 AAAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
 DB 591 AAAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 650  
 QY 541 ACCTGAGATGACCTGCTTTCAGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
 DB 651 ACCTGAGATGACCTGCTTTCAGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 710  
 QY 601 GTGCTAGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 626  
 DB 711 GTGCTAGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 736

## RESULT 11

US-08-394-388A-28

Sequence 28, Application US/08394388A

Patent No. 6753162

GENERAL INFORMATION:

APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/394,388A  
FILING DATE: 24-FEB-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1304 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-394-388A-28

Query Match 43.2%; Score 622.8; DB 4; Length 1304;  
Best Local Similarity 99.7%; Pred. No. 1.3e-137;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAACGGGGAGGCGCTTTTAAAGCACTGCTTCTGCTGCAAGCTGGGCTCCCTCCA 60  
DB 111 ATGAACGGGGAGGCGCTTTTAAAGCACTGCTTCTGCTGCAAGCTGGGCTCCCTCCA 170  
QY 61 GCAGCCACTCAGGAGGAGAAAGTGTCTGTGCGCAAAAAAGGGGATACAGTGAAGTGAAC 120  
DB 171 GCAGCCACTCAGGAGGAGAAAGTGTGTCTGTGCGCAAAAAAGGGGATACAGTGAAGTGAAC 230  
QY 121 TGTACAGCTTCCCAAGAAAGAGATCAATTCCACTGGAAGAACTCAACAGATTAAG 180  
DB 231 TGTACAGCTTCCCAAGAAAGAGATCAATTCCACTGGAAGAACTCAACAGATTAAG 290  
QY 181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCAATCCAAAGTGAATGATCGGCT 240  
DB 291 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCAATCCAAAGTGAATGATCGGCT 350  
QY 241 GACTCAAGAAAGCTTTGGGACCAAGAAACTTCCCTGATCATCAAGAAATCTTAAG 300  
DB 351 GACTCAAGAAAGCTTTGGGACCAAGAAACTTCCCTGATCATCAAGAAATCTTAAG 410

QY 301 ATGAACTCAGATCTTAACTTGTGAATGAGAGACCAAGAGAGAGGTGCATTG 360  
DB 411 ATGAACTCAGATCTTAACTTGTGAATGAGAGACCAAGAGAGAGGTGCATTG 470  
QY 361 CTAGGTGGGATTGCTGCGCAACTGTGACCCACCTGCTTCAAGGGGAGAGCTGACC 420  
DB 471 CTAGGTGGGATTGCTGCGCAACTGTGACCCACCTGCTTCAAGGGGAGAGCTGACC 530  
QY 421 CTGACCTTGAGAGCCCTCGTAGTAGCCCTCAGTGCATATGAGAGTCCAAAGGCT 480  
DB 531 CTGACCTTGAGAGCCCTCGTAGTAGCCCTCAGTGCATATGAGAGTCCAAAGGCT 590  
QY 481 AAAACATACAGGGGGGAGAGACCTCTCGTGTCTCAGCTGAGAGTCCAGATATGAGC 540  
DB 591 AAAACATACAGGGGGGAGAGACCTCTCGTGTCTCAGCTGAGAGTCCAGATATGAGC 650  
QY 541 ACCTGACATGACCTGCTTGGAGAACCAAGAGAGGTGAGATTAAGATCATGCTG 600  
DB 651 ACCTGACATGACCTGCTTGGAGAACCAAGAGAGGTGAGATTAAGATCATGCTG 710  
QY 601 GTGCTAGCTTTCAGAGGCTCCGG 626  
DB 711 GTGCTAGCTTTCAGAGGCTCCAG 736

RESULT 12  
US-08-417-495-2  
Sequence 2, Application US/08417495  
Patent No. 5843728  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian et al.  
TITLE OF INVENTION: Redirection of Cellular Immunity by  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
SOFTWARE: Wordperfect (Version 5.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,495  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/203,866  
FILING DATE:  
APPLICATION NUMBER: US/07/847,566  
FILING DATE:  
APPLICATION NUMBER: 07/665,961  
FILING DATE: March 7, 1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/119002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1389 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)



US-08-417-495-2

Query Match 43.2%; Score 622.8; DB 2; Length 1389;  
Best Local Similarity 99.7%; Pred. No. 1.3e-137;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTCGCACTGGCGCTCTCCCA 60
DB 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTCGCACTGGCGCTCTCCCA 60
QY 61 GCAGCCCTCAGGGAGAAAGTGTGTGTGGCAAAAAAGGGATACAGTGAATGACC 120
DB 61 GCAGCCCTCAGGGAGAAAGTGTGTGTGGCAAAAAAGGGATACAGTGAATGACC 120
QY 121 TGTACAGCTTCCGAGAGAGAGATCAATTCACATGGAATACTCCACAGATAAG 180
DB 121 TGTACAGCTTCCGAGAGAGAGATCAATTCACATGGAATACTCCACAGATAAG 180
QY 181 ATTCTGGAAATCAGGGCTCTCTTTTAATTAAGATCCTCAAGTGAATGACGCT 240
DB 181 ATTCTGGAAATCAGGGCTCTCTTTTAATTAAGATCCTCAAGTGAATGACGCT 240
QY 241 GACTCAAGAGAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 241 GACTCAAGAGAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
QY 301 ATGAAGACTCAGTACTTACATCTGTGAAGTGGAGCAAGAGAGAGAGTCAATTG 360
DB 301 ATGAAGACTCAGTACTTACATCTGTGAAGTGGAGCAAGAGAGAGAGTCAATTG 360
QY 361 CTAGTGTTCGATGACTGCACTGACACCCACCTGCTTCAAGGGGAGAGCTGACC 420
DB 361 CTAGTGTTCGATGACTGCACTGACACCCACCTGCTTCAAGGGGAGAGCTGACC 420
QY 421 CTGACCTTTGAGAGAGCCCCCTGTGTAAGCCCTCAGTCAATGAGAGTCAAGGGGT 480
DB 421 CTGACCTTTGAGAGAGCCCCCTGTGTAAGCCCTCAGTCAATGAGAGTCAAGGGGT 480
QY 481 AAAAACAATACAGGGGGGAGAGACCTCTGCTGTGTAAGTGGAGTCCAGATAGTGGC 540
DB 481 AAAAACAATACAGGGGGGAGAGACCTCTGCTGTGTAAGTGGAGTCCAGATAGTGGC 540
QY 541 ACCTGACATGACACTGTCTTGCAAGACAGAGAGAGTGAAGTCAAAATAGACATGCTG 600
DB 541 ACCTGACATGACACTGTCTTGCAAGACAGAGAGAGTGAAGTCAAAATAGACATGCTG 600
QY 601 GTGCTAGCTTTCCAGAGGCTCTCGG 626
DB 601 GTGCTAGCTTTCCAGAGGCTCTCGG 626
```

RESULT 13

US-08-284-391B-2

Sequence 2, Application US/08284391B

Patent No. 5851828

GENERAL INFORMATION:

APPLICANT: Seed, Brian

APPLICANT: Banepour, Babak

APPLICANT: Romeo, Charles

APPLICANT: Kolanus, Walidemar

TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED

TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESS: 176 Federal Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391B

FILING DATE: 02-AUG-1994

CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/195,395

FILING DATE: 14-FEB-1994

APPLICATION NUMBER: 07/847,566

FILING DATE: 06-MAR-1992

APPLICATION NUMBER: 07/665,961

FILING DATE: 07-MAR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Elbing, Karen L.

REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/247001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200

TELEFAX: 617-428-7045

TELEX:

INFORMATION FOR SRO ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1389 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-284-391B-2

Query Match 43.2%; Score 622.8; DB 2; Length 1389;  
Best Local Similarity 99.7%; Pred. No. 1.3e-137;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTCGCACTGGCGCTCTCCCA 60
DB 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTCGCACTGGCGCTCTCCCA 60
QY 61 GCAGCCCTCAGGGAGAAAGTGTGTGTGGCAAAAAAGGGATACAGTGAATGACC 120
DB 61 GCAGCCCTCAGGGAGAAAGTGTGTGTGGCAAAAAAGGGATACAGTGAATGACC 120
QY 121 TGTACAGCTTCCGAGAGAGAGATCAATTCACATGGAATACTCCACAGATAAG 180
DB 121 TGTACAGCTTCCGAGAGAGAGATCAATTCACATGGAATACTCCACAGATAAG 180
QY 181 ATTCTGGAAATCAGGGCTCTCTTTTAATTAAGATCCTCAAGTGAATGACGCT 240
DB 181 ATTCTGGAAATCAGGGCTCTCTTTTAATTAAGATCCTCAAGTGAATGACGCT 240
QY 241 GACTCAAGAGAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 241 GACTCAAGAGAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGCAAGAGAGAGTGAATGACC 360
DB 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGCAAGAGAGAGTGAATGACC 360
QY 361 CTAGTGTTCGATGACTGCACTGACACCCACCTGCTTCAAGGGGAGAGCTGACC 420
DB 361 CTAGTGTTCGATGACTGCACTGACACCCACCTGCTTCAAGGGGAGAGCTGACC 420
QY 421 CTGACCTTTGAGAGAGCCCCCTGTGTAAGCCCTCAGTCAATGAGAGTCAAGGGGT 480
DB 421 CTGACCTTTGAGAGAGCCCCCTGTGTAAGCCCTCAGTCAATGAGAGTCAAGGGGT 480
QY 481 AAAAACAATACAGGGGGGAGAGACCTCTGCTGTGTAAGTGGAGTCCAGATAGTGGC 540
DB 481 AAAAACAATACAGGGGGGAGAGACCTCTGCTGTGTAAGTGGAGTCCAGATAGTGGC 540
QY 541 ACCTGACATGACACTGTCTTGCAAGACAGAGAGAGTGAAGTCAAAATAGACATGCTG 600
DB 541 ACCTGACATGACACTGTCTTGCAAGACAGAGAGAGTGAAGTCAAAATAGACATGCTG 600
```



## ATTORNEY/AGENT INFORMATION:

NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:

## INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1389 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-394-388A-2

Query Match 43.2%; Score 622.8; DB 4; Length 1389;

Best Local Similarity 99.7%; Pred. No. 1,36-137;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGAACCGGGAGTCCCTTTTATAGCACTTGCTGCTGCAACTGGCGCTCTCCCA 60
Db 1 ATGAACCGGGAGTCCCTTTTATAGCACTTGCTGCTGCAACTGGCGCTCTCCCA 60
QY 61 GCAGCCACTCAGGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGAATGACC 120
Db 61 GCAGCCACTCAGGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGAATGACC 120
QY 121 TGTACAGCTTCCGAAAGAAAGAGCATACAAATTCACCTGGAATAAATCTCCAACTAGTAAG 180
Db 121 TGTACAGCTTCCGAAAGAAAGAGCATACAAATTCACCTGGAATAAATCTCCAACTAGTAAG 180
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGTCCATCAAGCTGAATGATCGGCT 240
Db 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGTCCATCAAGCTGAATGATCGGCT 240
QY 241 GACTCAGAAGAAAGCTTTGGGACCAAGAAATCTCCCGCTGATCATCAAGATCTTAAG 300
Db 241 GACTCAGAAGAAAGCTTTGGGACCAAGAAATCTCCCGCTGATCATCAAGATCTTAAG 300
QY 301 ATAGAAGACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAAAGAGGTGCAATTG 360
Db 301 ATAGAAGACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAAAGAGGTGCAATTG 360
QY 361 CTAGTGTTCGATGACTGCAACTCTGACCACTGACCACTGCTTCAGGGGCAAGGCTGACC 420
Db 361 CTAGTGTTCGATGACTGCAACTCTGACCACTGACCACTGCTTCAGGGGCAAGGCTGACC 420
QY 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTCAATGTAGGAGTCCAAAGGGGT 480
Db 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTCAATGTAGGAGTCCAAAGGGGT 480
QY 481 AAAAATATACAGGGGGGAGAACCTCTCCGTCTCAGCTGAGCTCAGATATAGGC 540
Db 481 AAAAATATACAGGGGGGAGAACCTCTCCGTCTCAGCTGAGCTCAGATATAGGC 540
QY 541 AACTGACATGCACTGTCTTCAAGAACCAAGAAAGGTGAGTTCAAATATGACATCGTG 600
Db 541 AACTGACATGCACTGTCTTCAAGAACCAAGAAAGGTGAGTTCAAATATGACATCGTG 600
QY 601 GTGCTAGCTTCCAGAGGCTCCGG 626
Db 601 GTGCTAGCTTCCAGAGGCTCCAG 626
```

Search completed: September 20, 2005, 10:22:34  
Job time : 300 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2005, 10:22:42 ; Search time 2458 Seconds

(without alignments)  
3903.215 Million cell updates/sec

Title: US-09-936-702-4  
Perfect score: 1440  
Sequence: 1 atgacacgggagtccttc.....gtcagcagcataactg 1440

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 3331285599 residues

Total number of hits satisfying chosen parameters: 1478644

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/1/pubpna/PCr_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
4:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
5:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
6:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
7:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
8:	/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12:	/cgn2_6/ptodata/1/pubpna/US09D_PUBCOMB.seq:*
13:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16:	/cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17:	/cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18:	/cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19:	/cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
20:	/cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
21:	/cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
22:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
23:	/cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
24:	/cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	625	43.4	8911	18	US-10-612-192-3
2	624.4	43.4	1377	14	US-10-103-597A-38
3	624.4	43.4	1377	15	US-10-188-444-38
4	624.4	43.4	1742	14	US-10-151-274-7
5	624.4	43.4	1742	18	US-10-641-643-1013
6	622.8	43.2	1304	10	US-09-939-537-28
7	622.8	43.2	1389	10	US-09-939-537-2

8	622.8	43.2	1389	11	US-09-243-008-2	Sequence 2, Appl1
9	622.8	43.2	1599	10	US-09-939-537-3	Sequence 3, Appl1
10	622.8	43.2	1599	11	US-09-243-008-3	Sequence 3, Appl1
11	622.8	43.2	1728	10	US-09-939-537-1	Sequence 1, Appl1
12	622.8	43.2	1728	11	US-09-243-008-1	Sequence 1, Appl1
13	622.8	43.2	1742	11	US-09-891-119A-8	Sequence 8, Appl1
14	622.8	43.2	1745	16	US-10-024-329-1	Sequence 1, Appl1
15	622.8	43.2	3084	15	US-10-207-655-169	Sequence 169, App
16	621.2	43.1	1273	11	US-09-891-119A-1	Sequence 1, Appl1
17	621.2	43.1	1416	15	US-10-157-408-2	Sequence 2, Appl1
18	621.2	43.1	1416	15	US-10-097-04A-2	Sequence 2, Appl1
19	621.2	43.1	1416	15	US-10-769-247-2	Sequence 2, Appl1
20	613.4	42.6	1796	8	US-08-485-163-2	Sequence 2, Appl1
21	613.4	42.6	1796	9	US-09-766-995-1	Sequence 1, Appl1
22	612.2	42.5	1149	8	US-08-485-163-6	Sequence 6, Appl1
23	612.2	42.5	1149	9	US-09-766-995-5	Sequence 5, Appl1
24	612.2	42.5	3273	17	US-10-397-565-7	Sequence 7, Appl1
25	612.2	42.5	11228	17	US-10-397-565-6	Sequence 6, Appl1
26	612	42.5	617	20	US-10-493-676-7	Sequence 7, Appl1
27	612	42.5	2482	8	US-08-485-163-4	Sequence 4, Appl1
28	612	42.5	2482	9	US-09-766-995-3	Sequence 3, Appl1
29	610	42.4	1350	20	US-10-493-676-10	Sequence 10, Appl1
30	610	42.4	1353	20	US-10-493-676-5	Sequence 5, Appl1
31	609	42.3	609	19	US-10-466-835-24	Sequence 24, Appl1
32	607.4	42.2	609	22	US-10-513-725-9	Sequence 9, Appl1
33	607.4	42.2	609	22	US-10-512-966-6	Sequence 6, Appl1
34	607.4	42.2	719	10	US-09-939-537-30	Sequence 30, Appl1
35	571.4	39.7	573	16	US-10-024-329-31	Sequence 31, Appl1
36	552.2	38.3	1508	15	US-10-157-408-5	Sequence 5, Appl1
37	552.2	38.3	1508	15	US-10-097-04A-5	Sequence 5, Appl1
38	552.2	38.3	1508	19	US-10-769-247-5	Sequence 5, Appl1
39	549.4	38.2	1113	9	US-09-759-841-5	Sequence 5, Appl1
40	542.4	37.7	563	14	US-10-794-774-24	Sequence 24, Appl1
41	542.4	37.7	563	21	US-10-073-118-25	Sequence 25, Appl1
42	541.8	37.6	2379	14	US-10-073-118-25	Sequence 25, Appl1
43	541.8	37.6	2379	21	US-10-794-774-25	Sequence 25, Appl1
44	535.6	37.2	1759	9	US-09-934-060A-12	Sequence 12, Appl1
45	535.6	37.2	2159	9	US-09-934-060A-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1  
US-10-612-192-3  
Sequence 3, Application US/10612192  
Publication No. US20040076636A1  
GENERAL INFORMATION:  
APPLICANT: Pal, Ranajit  
APPLICANT: Markham, Phillip  
APPLICANT: Keen, Timothy  
APPLICANT: Whitney, Stephen  
APPLICANT: Kalyanaraman, V.S.  
TITLE OR INVENTION: HIV Immunogenic Complexes  
FILE REFERENCE: 00711 CIP  
CURRENT APPLICATION NUMBER: US/10/612,192  
CURRENT FILING DATE: 2003-07-02  
PRIOR APPLICATION NUMBER: US 09/905,962  
PRIOR FILING DATE: 2001-07-17  
PRIOR APPLICATION NUMBER: US 09/479,675  
PRIOR FILING DATE: 2000-01-07  
PRIOR APPLICATION NUMBER: US 09/075,544  
PRIOR FILING DATE: 1998-05-11  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 3  
LENGTH: 8911  
TYPE: DNA  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: Expression Vector PTK13+Neo4  
US-10-612-192-3

Query Match	Similarity	43.4%	Score 625	DB 18	Length 8911
Best Local	Similarity	100.0%	Pred. No. 8.2e-171		
Matches	625	Conservative	0	Mismatches	0
				Indels	Gaps
Qy	1	ATGAACCGGGAGTCCCTTTTAGGCACTTCTCTGATGTCGACACTGGCGCTCCTCCCA	60		
Db	3815	ATGAACCGGGAGTCCCTTTTAGGCACTTCTCTGATGTCGACACTGGCGCTCCTCCCA	3874		
Qy	61	GCAGCCACTCAGGGAAAGAAAGTGGTCTGGGCAAAAAAGGGGATACAGTGAACCTGAC	120		
Db	3875	GCAGCCACTCAGGGAAAGAAAGTGGTCTGGGCAAAAAAGGGGATACAGTGAACCTGAC	3934		
Qy	121	TGTACAGCTTCCCAAGAAAGAGCATCAATTCCAGTGGAAAAATCCCAACAGATAAG	180		
Db	3935	TGTACAGCTTCCCAAGAAAGAGCATCAATTCCAGTGGAAAAATCCCAACAGATAAG	3999		
Qy	181	ATTCTGGAAATCAGGGCTCTCTTTTAACTAAAGGTCATCCAGCTGAATGATCGCT	240		
Db	3995	ATTCTGGAAATCAGGGCTCTCTTTTAACTAAAGGTCATCCAGCTGAATGATCGCT	405		
Qy	241	GACTCAAGAAAGCTTTTGGGACCAAGAAACTTCCCCCTGATCATCAAGATTCTTAAG	300		
Db	4055	GACTCAAGAAAGCTTTTGGGACCAAGAAACTTCCCCCTGATCATCAAGATTCTTAAG	4114		
Qy	301	ATAGAGACTGAGACTTCACTCTGTAAAGTGGAGGACCAAGAGAGAGGTGCATTTG	360		
Db	4115	ATAGAGACTGAGACTTCACTCTGTAAAGTGGAGGACCAAGAGAGAGGTGCATTTG	4174		
Qy	361	CTAGTGTTCCGATTGACTGCCCACTCTGACACCCACTGCTTTCAGGGGACAGCCTGACC	420		
Db	4175	CTAGTGTTCCGATTGACTGCCCACTCTGACACCCACTGCTTTCAGGGGACAGCCTGACC	4233		
Qy	421	CTGACCTTGGAGAGCCCCCTGTGTAGTACGCCCTCAGTGCATGTAGAGTCCAAAGGGT	480		
Db	4235	CTGACCTTGGAGAGCCCCCTGTGTAGTACGCCCTCAGTGCATGTAGAGTCCAAAGGGT	4294		
Qy	481	AAAAACATACAGGGGGGGAAGAACCTCTCCGCTGTCACTGGAGCTCAAGATGTGAGC	540		
Db	4295	AAAAACATACAGGGGGGGAAGAACCTCTCCGCTGTCACTGGAGCTCAAGATGTGAGC	4354		
Qy	541	ACCTGACATCACTGTCTTTCGAGAACCAAGAAAGTGGAGTTCAAAATAGACATCGTG	600		
Db	4355	ACCTGACATCACTGTCTTTCGAGAACCAAGAAAGTGGAGTTCAAAATAGACATCGTG	4414		
Qy	601	GTGCTAGCTTTTCAGAGGCTCCG	625		
Db	4415	GTGCTAGCTTTTCAGAGGCTCCG	4439		
RESULT 2					
US-10-103-597A-38					
: Sequence 38, Application US/10103597A					
: Publication No. US20030096432A1					
: GENERAL INFORMATION:					
: APPLICANT: Jakobsen, Bent Karsten					
: TITLE OF INVENTION: Screening Methods					
: FILE REFERENCE: 102286.142					
: CURRENT APPLICATION NUMBER: US/10/103.597A					
: CURRENT FILING DATE: 2002-10-17					
: PRIOR APPLICATION NUMBER: PCT/GB00/03579					
: PRIOR FILING DATE: 2000-09-18					
: PRIOR APPLICATION NUMBER: GB 9922352.1					
: PRIOR FILING DATE: 1999-09-21					
: NUMBER OF SEQ ID NOS: 39					
: SOFTWARE: FastSeq for Windows Version 4.0					
: SEQ ID NO 38					
: LENGTH: 1377					
: TYPE: DNA					
: ORGANISM: Homo sapiens					
: FEATURE:					
: NAME/KEY: CDS					
: LOCATION: (1)...(1377)					
: US-10-103-597A-38					

Query Match	Similarity	43.4%	Score 624.4	DB 14	Length 1377
Best Local	Similarity	99.8%	Pred. No. 6.6e-171		
Matches	625	Conservative	0	Mismatches	1
				Indels	0
				Gaps	0
Qy	1	ATGAACCGGGGAGTCCCTTTTAGGCACTTGTGCTGCTGCAATGCGCTCTCTCCA	60		
Db	1	ATGAACCGGGGAGTCCCTTTTAGGCACTTGTGCTGCTGCAATGCGCTCTCTCCA	60		
Qy	61	GCACCACTCAGGGGAAAGAAAGTGTGTGGGCAAAAAGGGGATACAGTGAAC	120		
Db	61	GCACCACTCAGGGGAAAGAAAGTGTGTGGGCAAAAAGGGGATACAGTGAAC	120		
Qy	121	TGTACAGCTTCCAGAAAGAGAGATCAATTCACATGAAAAAATCCCAACGATTAAG	180		
Db	121	TGTACAGCTTCCAGAAAGAGAGATCAATTCACATGAAAAAATCCCAACGATTAAG	180		
Qy	181	ATTCTGGAAATCAGGGGCTCTCTTTTAACTTAAGGTCATCAAGCTGAATGTCGCT	240		
Db	181	ATTCTGGAAATCAGGGGCTCTCTTTTAACTTAAGGTCATCAAGCTGAATGTCGCT	240		
Qy	241	GACTCAAGAAAGACCTTTGGGACCAAGGAACTTCCCTCGATCATCAAGATCTTAAG	300		
Db	241	GACTCAAGAAAGACCTTTGGGACCAAGGAACTTCCCTCGATCATCAAGATCTTAAG	300		
Qy	301	ATAGAAGCTCAGATTAATCAATCTGTGAATGAGGACCAAGAAAGAGAGTCAATTG	360		
Db	301	ATAGAAGCTCAGATTAATCAATCTGTGAATGAGGACCAAGAAAGAGAGTCAATTG	360		
Qy	361	CTAGTGTTCGAGTTGACTGCCAATCTGACACCACTGCTTCAAGGGCAGAGCTTGAC	420		
Db	361	CTAGTGTTCGAGTTGACTGCCAATCTGACACCACTGCTTCAAGGGCAGAGCTTGAC	420		
Qy	421	CTGACCTTGGAGACCCCTGTGAGTAAAGCCCTCAGTCAATGAAGAGTCAAGGGGT	480		
Db	421	CTGACCTTGGAGACCCCTGTGAGTAAAGCCCTCAGTCAATGAAGAGTCAAGGGGT	480		
Qy	481	AAAAACATACAGGGGGGAGAGACCTCTGCTGCTGCAAGTGAAGCTCAGAGTAGTGGC	540		
Db	481	AAAAACATACAGGGGGGAGAGACCTCTGCTGCTGCAAGTGAAGCTCAGAGTAGTGGC	540		
Qy	541	ACCTGAGATGCACTGTCTTTCAGAACCAAGAAAGTGAATCAAAATGACATGCTG	600		
Db	541	ACCTGAGATGCACTGTCTTTCAGAACCAAGAAAGTGAATCAAAATGACATGCTG	600		
Qy	601	GTGCTAGCTTCCAGAAAGGCTCCGG	626		
Db	601	GTGCTAGCTTCCAGAAAGGCTCCAG	626		
RESULT 3					
US-10-188-444-38					
; Sequence 38, Application US/10188444					
; Publication No. US20030104635A1					
; GENERAL INFORMATION:					
; APPLICANT: Jakobsen, Bent Karsten					
; TITLE OF INVENTION: Screening Methods					
; FILE REFERENCE: 102286.142 (CIP)					
; CURRENT APPLICATION NUMBER: US/10/188.444					
; CURRENT FILING DATE: 2002-07-02					
; PRIOR APPLICATION NUMBER: PCT/G800/03579					
; PRIOR FILING DATE: 2000-09-18					
; PRIOR APPLICATION NUMBER: GB 9922352.1					
; PRIOR FILING DATE: 1999-09-21					
; NUMBER OF SEQ ID NOS: 39					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 38					
; LENGTH: 1377					
; TYPE: DNA					
; ORGANISM: Homo sapiens					
; FEATURE:					
; NAME/KEY: CDS					
; LOCATION: (1)...(1377)					

US-10-188-444-38

43.4%; Score 624.4; DB 15; Length 1377;

Query Match Best Local Similarity 99.8%; Pred. No. 6.6e-111; Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAACCGGGAGATCCCTTTTAGGCACTTCTGTGTGCACTGAGCGCTCTCCCA 60
DB 1 ATGAACCGGGAGATCCCTTTTAGGCACTTCTGTGTGCACTGAGCGCTCTCCCA 60
QY 61 GCAGCCACTCAGGAAAGAAAGTGTCTGGGCAAAAAGGGATACATGAACTGACC 120
DB 61 GCAGCCACTCAGGAAAGAAAGTGTCTGGGCAAAAAGGGATACATGAACTGACC 120
QY 121 TGTACAGCTTCCAGAAAGAGACATCAATTCACATGAAACCTCAACCGATTAAG 180
DB 121 TGTACAGCTTCCAGAAAGAGACATCAATTCACATGAAACCTCAACCGATTAAG 180
QY 181 ATTCTGGAAATCAGGGCTCTCTTAATTAAGTCCATCAAGCTGAATGATCGGCT 240
DB 181 ATTCTGGAAATCAGGGCTCTCTTAATTAAGTCCATCAAGCTGAATGATCGGCT 240
QY 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
QY 301 ATGAAGACTCAGATTAATCTGTGAAGTGAAGGACCAAGAGGAGGAGTCAATTG 360
DB 301 ATGAAGACTCAGATTAATCTGTGAAGTGAAGGACCAAGAGGAGGAGTCAATTG 360
QY 361 CTAGTGTGGATTTAGCTGCAACTGACACCCAGCTGCTTCAAGGGGAGAGCTGACC 420
DB 361 CTAGTGTGGATTTAGCTGCAACTGACACCCAGCTGCTTCAAGGGGAGAGCTGACC 420
QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCACTGATGATGAGTCAAGGGGT 480
DB 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCACTGATGATGAGTCAAGGGGT 480
QY 481 AAAAACAATACAGGGGGGAGAACCTCTCGTGTCTGACGTGAGCTCCAGAGATGAGGC 540
DB 481 AAAAACAATACAGGGGGGAGAACCTCTCGTGTCTGACGTGAGCTCCAGAGATGAGGC 540
QY 541 ACCTGACATGACACTGTCTTGCAGAACCAAGAAAGGTGAGTCAAAATAGACATGCTG 600
DB 541 ACCTGACATGACACTGTCTTGCAGAACCAAGAAAGGTGAGTCAAAATAGACATGCTG 600
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
DB 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
```

RESULT 4

```
US-10-151-274-7
; Sequence 7, Application US/10151274
; Publication No. US20030064071A1
; GENERAL INFORMATION:
; APPLICANT: Liltman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geljtenbeck, Tneo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY
; TITLE OF INVENTION: INTO
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/10/151.274
; PRIOR APPLICATION NUMBER: 2002-05-20
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1742
; TYPE: DNA
```

; ORGANISM: Homo sapiens

US-10-151-274-7

43.4%; Score 624.4; DB 14; Length 1742;

Query Match Best Local Similarity 99.8%; Pred. No. 7.1e-171; Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAACCGGGAGATCCCTTTTAGGCACTTCTGTGTGCACTGAGCGCTCTCCCA 60
DB 76 ATGAACCGGGAGATCCCTTTTAGGCACTTCTGTGTGCACTGAGCGCTCTCCCA 135
QY 61 GCAGCCACTCAGGAAAGAAAGTGTCTGGGCAAAAAGGGATACATGAACTGACC 120
DB 136 GCAGCCACTCAGGAAAGAAAGTGTCTGGGCAAAAAGGGATACATGAACTGACC 195
QY 121 TGTACAGCTTCCAGAAAGAGACATCAATTCACATGAAACCTCAACCGATTAAG 180
DB 196 TGTACAGCTTCCAGAAAGAGACATCAATTCACATGAAACCTCAACCGATTAAG 255
QY 181 ATTCTGGAAATCAGGGCTCTCTTAATTAAGTCCATCAAGCTGAATGATCGGCT 240
DB 256 ATTCTGGAAATCAGGGCTCTCTTAATTAAGTCCATCAAGCTGAATGATCGGCT 315
QY 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 316 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 375
QY 301 ATGAAGACTCAGATTAATCTGTGAAGTGAAGGACCAAGAGGAGGAGTCAATTG 360
DB 376 ATGAAGACTCAGATTAATCTGTGAAGTGAAGGACCAAGAGGAGGAGTCAATTG 435
QY 361 CTAGTGTGGATTTAGCTGCAACTGACACCCAGCTGCTTCAAGGGGAGAGCTGACC 420
DB 436 CTAGTGTGGATTTAGCTGCAACTGACACCCAGCTGCTTCAAGGGGAGAGCTGACC 495
QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCACTGATGATGAGTCAAGGGGT 480
DB 496 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCACTGATGATGAGTCAAGGGGT 555
QY 481 AAAAACAATACAGGGGGGAGAACCTCTCGTGTCTGACGTGAGCTCCAGAGATGAGGC 540
DB 556 AAAAACAATACAGGGGGGAGAACCTCTCGTGTCTGACGTGAGCTCCAGAGATGAGGC 615
QY 541 ACCTGACATGACACTGTCTTGCAGAACCAAGAAAGGTGAGTCAAAATAGACATGCTG 600
DB 616 ACCTGACATGACACTGTCTTGCAGAACCAAGAAAGGTGAGTCAAAATAGACATGCTG 675
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
DB 676 GTGCTAGCTTTCCAGAAAGGCTCCGG 701
```

RESULT 5

```
US-10-641-643-1013
; Sequence 1013, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Sellhammer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; GENE EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US//10/641,643  
FILING DATE: 14-Aug-2003  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1013:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1742 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: g179143  
SEQUENCE DESCRIPTION: SEQ ID NO: 1013 :  
US-10-641-643-1013

Query Match 43.4%; Score 624.4; DB 18; Length 1742;  
Best Local Similarity 99.8%; Pred. No. 7.1e-171;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGCTGCTGCACTGGCGCTCTCCCA 60  
DB 76 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGCTGCTGCACTGGCGCTCTCCCA 135  
QY 61 GCAGCCACTAGGAGAAAGAGTGTCTGGGCAAAAAGGGATACAGTGAACAGACC 120  
DB 136 GCAGCCACTAGGAGAAAGAGTGTCTGGGCAAAAAGGGATACAGTGAACAGACC 195  
QY 121 TGTACGCTTCCAGAGAGAGATCAATTCACAGTGAAGAACTCCACAGATTAAG 180  
DB 196 TGTACGCTTCCAGAGAGAGATCAATTCACAGTGAAGAACTCCACAGATTAAG 255  
QY 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTTCATCAAGCTGAATCGCGCT 240  
DB 256 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTTCATCAAGCTGAATCGCGCT 315  
QY 241 GACTCAAGAAAGACCTTTGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 316 GACTCAAGAAAGACCTTTGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 375  
QY 301 ATGAAGACTCAGATTAATCATCTGTGAAGTGAAGAGAGAGAGAGAGTGAATTG 360  
DB 376 ATGAAGACTCAGATTAATCATCTGTGAAGTGAAGAGAGAGAGAGAGTGAATTG 435  
QY 361 CTAGTGTGGATTAAGTCTGCACTGCAACCACTGCTTCAAGGGGAGAGCTGACC 420  
DB 436 CTAGTGTGGATTAAGTCTGCACTGCAACCACTGCTTCAAGGGGAGAGCTGACC 495  
QY 421 CTGACCTTGAAGAGCCCTTGGTGAAGCCCTCACTGCAAGTGAAGATCCAGAGGCT 480  
DB 496 CTGACCTTGAAGAGCCCTTGGTGAAGCCCTCACTGCAAGTGAAGATCCAGAGGCT 555  
QY 481 AAAAACAATACAGGGGGAGAGACCTCTCCGTGTCTCAGCTGAGAGCTCCAGATTAAGTGC 540  
DB 556 AAAAACAATACAGGGGGAGAGACCTCTCCGTGTCTCAGCTGAGAGCTCCAGATTAAGTGC 615  
QY 541 ACCTGACATGACATGCTTTGCAAGAACAGAGAGAGTGAAGTTCAAATATGACATCTGTG 600  
DB 616 ACCTGACATGACATGCTTTGCAAGAACAGAGAGAGTGAAGTTCAAATATGACATCTGTG 675  
QY 601 GTCTAGCTTTCCAGAGAGGCTCCGG 626

DB 676 GTCTAGCTTTCCAGAGAGGCTCCGG 701

RESULT 6  
US-09-939-537-28  
Sequence 28, Application US/09939537  
Publication No. US20030138410A1  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
Ranapur, Babak  
Romeo, Charles  
Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,537  
FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1304 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 28:  
US-09-939-537-28

Query Match 43.2%; Score 622.8; DB 10; Length 1304;  
Best Local Similarity 99.7%; Pred. No. 1.9e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGCTGCTGCACTGGCGCTCTCCCA 60  
DB 111 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGCTGCTGCACTGGCGCTCTCCCA 170  
QY 61 GCAGCCACTAGGAGAAAGAGTGTCTGGGCAAAAAGGGATACAGTGAACAGACC 120  
DB 171 GCAGCCACTAGGAGAAAGAGTGTCTGGGCAAAAAGGGATACAGTGAACAGACC 230  
QY 121 TGTACGCTTCCAGAGAGAGATCAATTCACAGTGAAGAACTCCACAGATTAAG 180



Db 231 TGTACAGCTTCCAGAGAGACATCAATTCACCTGAGAAAACTCCAAACAGATAAG 230  
QY 181 ATTCTGGGAATCAGGGCTCTCTTCTTAATAAGGTCCATCAAGTGAATGCGGCT 240  
Db 291 ATTCTGGGAATCAGGGCTCTCTTCTTAATAAGGTCCATCAAGTGAATGCGGCT 350  
QY 241 GACTCAAGAGAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTTAG 300  
Db 351 GACTCAAGAGAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTTAG 410  
QY 301 ATGAGAGACTCAATATCTTATCTGTGAGAGTGAAGACCAAGAGAGAGTGCATTG 360  
Db 411 ATGAGAGACTCAATATCTTATCTGTGAGAGTGAAGACCAAGAGAGAGTGCATTG 470  
QY 361 CTAGTGTGGGATTTGACTGCAACTCTGACACCACTGCTTCAAGGGGAGAGCTGACC 420  
Db 471 CTAGTGTGGGATTTGACTGCAACTCTGACACCACTGCTTCAAGGGGAGAGCTGACC 530  
QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAGGGGT 480  
Db 531 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAGGGGT 590  
QY 481 AAAAACAATACAGGGGGGAGAGACCTCTCTGTGTCTGAGCTGCAAGATAGTGGC 540  
Db 591 AAAAACAATACAGGGGGGAGAGACCTCTCTGTGTCTGAGCTGCAAGATAGTGGC 650  
QY 541 ACCTGACATGACCTGTCTTGCAGAACCAAGAGAGTGAAGTTCAAAATAGACATGTC 600  
Db 651 ACCTGACATGACCTGTCTTGCAGAACCAAGAGAGTGAAGTTCAAAATAGACATGTC 710  
QY 601 GTGCTAGCTTTCCAGAGGCTCCGG 626  
Db 711 GTGCTAGCTTTCCAGAGGCTCCAG 736

RESULT 7  
US-09-939-537-2

; Sequence 2, Application US/09939537  
; Publication No. US20030138410A1

## GENERAL INFORMATION:

APPLICANT: Seed, Brian

Banapour, Babak

Romeo, Charles

Kolanus, Waldemar

TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INJECTED

CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

STREET: 176 Federal Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: IBM Compatible

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/939, 537

FILING DATE: 24-Aug-2001

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/284,391

FILING DATE: 02-AUG-1994

APPLICATION NUMBER: 08/195,395

FILING DATE: 14-FEB-1994

APPLICATION NUMBER: 07/847,566

FILING DATE: 06-MAR-1992

APPLICATION NUMBER: 07/665,961

FILING DATE: 07-MAR-1991

ATTORNEY/AGENT INFORMATION:

/

NAME: Elbing, Karen L  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1389 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-939-537-2

Query Match 43.2%; Score 622.8; DB 10; Length 1389;  
Best Local Similarity 99.7%; Pred. No. 1.9e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGACCGGGAGTCCCTTTTAGCACTGCTCTGTGCTGCACTGGGCGCTCCGCCA 60  
Db 1 ATGACCGGGAGTCCCTTTTAGCACTGCTCTGTGCTGCACTGGGCGCTCCGCCA 60  
QY 61 GCAGCACTCAGGGGAGAGAGTGTGTGGCAAAAAAGGGGATACAGTGAATGACC 120  
Db 61 GCAGCACTCAGGGGAGAGAGTGTGTGGCAAAAAAGGGGATACAGTGAATGACC 120  
QY 121 TGTACAGCTTCCAGAGAGAGACATACATTCCTGAGAAAACTCCACAGATTAAG 180  
Db 121 TGTACAGCTTCCAGAGAGAGACATACATTCCTGAGAAAACTCCACAGATTAAG 180  
QY 181 ATTCTGGGAATCAGGGCTCTCTTCTTAATAAGGTCCATCAAGTGAATGCGGCT 240  
Db 181 ATTCTGGGAATCAGGGCTCTCTTCTTAATAAGGTCCATCAAGTGAATGCGGCT 240  
QY 241 GACTCAAGAGAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTTAG 300  
Db 241 GACTCAAGAGAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTTAG 300  
QY 301 ATGAGAGACTCAATATCTTATCTGTGAGAGTGAAGACCAAGAGAGTGCATTG 360  
Db 301 ATGAGAGACTCAATATCTTATCTGTGAGAGTGAAGACCAAGAGAGTGCATTG 360  
QY 361 CTAGTGTGGGATTTGACTGCAACTCTGACACCACTGCTTCAAGGGGAGAGCTGACC 420  
Db 361 CTAGTGTGGGATTTGACTGCAACTCTGACACCACTGCTTCAAGGGGAGAGCTGACC 420  
QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAGGGGT 480  
Db 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAGGGGT 480  
QY 481 AAAAACAATACAGGGGGGAGAGACCTCTCTGTGTCTGAGCTGCAAGATAGTGGC 540  
Db 481 AAAAACAATACAGGGGGGAGAGACCTCTCTGTGTCTGAGCTGCAAGATAGTGGC 540  
QY 541 ACCTGACATGACCTGTCTTGCAGAACCAAGAGAGTGAAGTTCAAAATAGACATGTC 600  
Db 541 ACCTGACATGACCTGTCTTGCAGAACCAAGAGAGTGAAGTTCAAAATAGACATGTC 600  
QY 601 GTGCTAGCTTTCCAGAGGCTCCGG 626  
Db 601 GTGCTAGCTTTCCAGAGGCTCCAG 626

RESULT 8  
US-09-243-008-2

; Sequence 2, Application US/09243008  
; Publication No. US20040005334A1

GENERAL INFORMATION:

APPLICANT: Seed, Brian et al.

TITLE OF INVENTION: Redirection of Cellular Immunity by

Receptor Chimeras  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
OPERATING SYSTEM: IBM P.C. DOS (version 3.30)  
SOFTWARE: Wordperfect (version 5.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/243,008  
FILING DATE: 02-Feb-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/394,176  
FILING DATE: SEPTEMBER 11, 1995  
APPLICATION NUMBER: 08/203,866  
FILING DATE: February 28, 1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: March 6, 1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: March 7, 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Karen F. Lech, Ph.D  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/270001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1389 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-243-008-2

Query Match 43.2%; Score 622.8; DB 11; Length 1389;  
Best Local Similarity 99.7%; Pred. No. 1.9e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGGAGTCTCTTTTAAAGCACTTCTGTGCTGCAACTGGCGCTCTCCCA 60  
DB 1 ATGAACCGGGGAGTCTCTTTTAAAGCACTTCTGTGCTGCAACTGGCGCTCTCCCA 60  
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAAGTACC 120  
DB 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAAGTACC 120  
QY 121 TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGGAATACTCCACCAAGATAAG 180  
DB 121 TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGGAATACTCCACCAAGATAAG 180  
QY 181 ATCTGGGAATCAGGGCTCTCTTTAACTAAAGGTCCATCCAAAGCTGAATGTCGGCT 240  
DB 181 ATCTGGGAATCAGGGCTCTCTTTAACTAAAGGTCCATCCAAAGCTGAATGTCGGCT 240  
QY 241 GACTCAAGAAAGCCTTTGGGACCAAGAAAGTCCCGGTGATCATCAAGAAATCTTAAG 300  
DB 241 GACTCAAGAAAGCCTTTGGGACCAAGAAAGTCCCGGTGATCATCAAGAAATCTTAAG 300  
QY 301 ATGAAGACTCAGATACCTTACATCTGTGAAGTGAAGACCAAGAAAGAGGTGCAATTG 360  
DB 301 ATGAAGACTCAGATACCTTACATCTGTGAAGTGAAGACCAAGAAAGAGGTGCAATTG 360  
QY 361 CTAGTGTGGAGATTGACTGCCAACTGCAACCCACTGCTTCAAGGGGCGAGAGCTTGACC 420

DB 361 CTAGTGTGGAGATTGACTGCCAACTGCAACCCACTGCTTCAAGGGGCGAGAGCTTGACC 420  
QY 421 CTGACCTTGAAGAGCCCCCTGGTAGTAGCCCTCAGTGCATATGTAGAGTCAAGAGGT 480  
DB 421 CTGACCTTGAAGAGCCCCCTGGTAGTAGCCCTCAGTGCATATGTAGAGTCAAGAGGT 480  
QY 481 AAAACATACAGGGGGGGAAGACCTCTCGTGTCTCAGCTGAGCTCCAGATAGTGC 540  
DB 481 AAAACATACAGGGGGGGAAGACCTCTCGTGTCTCAGCTGAGCTCCAGATAGTGC 540  
QY 541 ACCTGACATGCACTCTCTTGAAGAACCAAGAGGTGAGTCAATATGACATCTG 600  
DB 541 ACCTGACATGCACTCTCTTGAAGAACCAAGAGGTGAGTCAATATGACATCTG 600  
QY 601 GTGCTAGCTTTCAGAAAGCCTCCGG 626  
DB 601 GTGCTAGCTTTCAGAAAGCCTCCAG 626

RESULT 9  
US-09-939-537-3  
Sequence 3, Application US/09939537  
Publication No. US20030138410A1  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
Banapour, Babak  
Romeo, Charles  
Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,537  
FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1599 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-939-537-3

Query Match 43.2%; Score 622.8; DB 10; Length 1599;  
 Best Local Similarity 99.7%; Pred. No. 2e-170;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGACCGGGGAGTCCCTTTTATGACACTTCTGTGTGTCGAATGCGCTCTCCCA 60  
 Db 1 ATGACCGGGGAGTCCCTTTTATGACACTTCTGTGTGTCGAATGCGCTCTCCCA 60  
 QY 61 GCAGCCACTCAGGGAAGAAAGTGTCTGTGGCAAAAAGGGGATACAGTGAATGACC 120  
 Db 61 GCAGCCACTCAGGGAAGAAAGTGTCTGTGGCAAAAAGGGGATACAGTGAATGACC 120  
 QY 121 TGTACAGCTTCCAGAGAAAGATCAATTCACGTAAGAACTCCACGATTAAG 180  
 Db 121 TGTACAGCTTCCAGAGAAAGATCAATTCACGTAAGAACTCCACGATTAAG 180  
 QY 181 ATTCTGGGAATCAGGGCTCTCTTTAATTAAGTCCATCAAGTGAATGATCGGCT 240  
 Db 181 ATTCTGGGAATCAGGGCTCTCTTTAATTAAGTCCATCAAGTGAATGATCGGCT 240  
 QY 241 GACTCAAGAAAGACCTTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
 Db 241 GACTCAAGAAAGACCTTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
 QY 301 ATGAGAGACTCAGATCTTACATCTGTGAAGTGGAGACCAAGAGAGGATGCAATG 360  
 Db 301 ATGAGAGACTCAGATCTTACATCTGTGAAGTGGAGACCAAGAGAGGATGCAATG 360  
 QY 361 CTAGTGTCCGATTTAGTGCCTCACTGACACCCAGCTGTTCAAGGGGAGAGCTGACC 420  
 Db 361 CTAGTGTCCGATTTAGTGCCTCACTGACACCCAGCTGTTCAAGGGGAGAGCTGACC 420  
 QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATGTAGAGTCCAAAGGGGT 480  
 Db 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATGTAGAGTCCAAAGGGGT 480  
 QY 481 AAAAACAATAAGGGGGGAGAACCTCTCGTGTCTGAGTGCAGTCCAGATAGTGGC 540  
 Db 481 AAAAACAATAAGGGGGGAGAACCTCTCGTGTCTGAGTGCAGTCCAGATAGTGGC 540  
 QY 541 ACCTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 600  
 Db 541 ACCTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 600  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTTCGG 626  
 Db 601 GTGCTAGCTTTCCAGAAAGGCTTCAG 626

RESULT 10  
 US-09-243-008-3  
 Sequence 3, Application US/09243008  
 Publication No. US2004005334A1

GENERAL INFORMATION:

APPLICANT: Seed, Brian et al.  
 TITLE OF INVENTION: Redirection of Cellular Immunity by  
 Receptor Chimeras

NUMBER OF SEQUENCES: 40  
 CORRESPONDENCE ADDRESSES:

ADDRESSEE: Fish & Richardson P.C.  
 STREET: 225 Franklin Street  
 CITY: Boston

STATE: MA

COUNTRY: USA  
 ZIP: 02110-2804

COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM PS/2 Model 502 or 55SX  
 OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)

SOFTWARE: Wordperfect (Version 5.0)  
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/243,008  
 FILING DATE: 02-Feb-1999

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/394,176

FILING DATE: SEPTEMBER 11, 1995

APPLICATION NUMBER: 08/203,866

FILING DATE: February 28, 1994

APPLICATION NUMBER: 07/847,566

FILING DATE: March 6, 1992

APPLICATION NUMBER: 07/665,961

FILING DATE: March 7, 1991

ATTORNEY/AGENT INFORMATION:

NAME: Karen F. Lech, Ph.D.

REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00766/270001

TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1599 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-243-008-3  
 Query Match 43.2%; Score 622.8; DB 11; Length 1599;  
 Best Local Similarity 99.7%; Pred. No. 2e-170;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGACCGGGGAGTCCCTTTTATGACACTTCTGTGTGTCGAATGCGGCTCTCCCA 60  
 Db 1 ATGACCGGGGAGTCCCTTTTATGACACTTCTGTGTGTCGAATGCGGCTCTCCCA 60  
 QY 61 GCAGCCACTCAGGGAAGAAAGTGTCTGTGGCAAAAAGGGGATACAGTGAATGACC 120  
 Db 61 GCAGCCACTCAGGGAAGAAAGTGTCTGTGGCAAAAAGGGGATACAGTGAATGACC 120  
 QY 121 TGTACAGCTTCCAGAGAAAGATCAATTCACGTAAGAACTCCACGATTAAG 180  
 Db 121 TGTACAGCTTCCAGAGAAAGATCAATTCACGTAAGAACTCCACGATTAAG 180  
 QY 181 ATTCTGGGAATCAGGGCTCTCTTTAATTAAGTCCATCAAGTGAATGATCGGCT 240  
 Db 181 ATTCTGGGAATCAGGGCTCTCTTTAATTAAGTCCATCAAGTGAATGATCGGCT 240  
 QY 241 GACTCAAGAAAGACCTTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
 Db 241 GACTCAAGAAAGACCTTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
 QY 301 ATGAGAGACTCAGATCTTACATCTGTGAAGTGGAGACCAAGAGAGGATGCAATG 360  
 Db 301 ATGAGAGACTCAGATCTTACATCTGTGAAGTGGAGACCAAGAGAGGATGCAATG 360  
 QY 361 CTAGTGTCCGATTTAGTGCCTCACTGACACCCAGCTGTTCAAGGGGAGAGCTGACC 420  
 Db 361 CTAGTGTCCGATTTAGTGCCTCACTGACACCCAGCTGTTCAAGGGGAGAGCTGACC 420  
 QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATGTAGAGTCCAAAGGGGT 480  
 Db 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATGTAGAGTCCAAAGGGGT 480  
 QY 481 AAAAACAATAAGGGGGGAGAACCTCTCGTGTCTCAGTGAAGTCCAGATAGTGGC 540  
 Db 481 AAAAACAATAAGGGGGGAGAACCTCTCGTGTCTCAGTGAAGTCCAGATAGTGGC 540  
 QY 541 ACCTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 600  
 Db 541 ACCTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 600

QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
DB 601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626

## RESULT 11

US-09-939-537-1  
Sequence 1, Application US/09939537  
Publication No. US20030138410A1  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
Banapour, Babak  
Romeo, Charles  
Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,537  
FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1728 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-939-537-1

Query Match 43.2%; Score 622.8; DB 10; Length 1728;  
Best Local Similarity 99.7%; Pred. No. 2.1e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGACCGGGAGGCCCTTTAGCACTTGTGCTGCAACTGGGCTCCGCCA 60  
DB 1 ATGACCGGGAGGCCCTTTAGCACTTGTGCTGCAACTGGGCTCCGCCA 60  
QY 61 GCAGCCACTCAGGAAAGAAAGTGTGCTGCGCAAAAAAGGGGATACAGTGAACCTGACC 120  
DB 61 GCAGCCACTCAGGAAAGAAAGTGTGCTGCGCAAAAAAGGGGATACAGTGAACCTGACC 120  
QY 121 TGTACAGCTTCCAGAAAGAGCATCAATTCCACTGGAATACTCAACAGATTAAG 180

DB 121 TGTACAGCTTCCAGAAAGAGCATCAATTCCACTGGAATACTCAACAGATTAAG 180  
QY 181 ATTCTGGAAATCAGGGCTCTCTTAACTAAAGGTCATCCAAAGTGAATGCGCGT 240  
DB 181 ATTCTGGAAATCAGGGCTCTCTTAACTAAAGGTCATCCAAAGTGAATGCGCGT 240  
QY 241 GACTCAAGAAAGACCTTTGGAGCAAGAAACCTTCCCTGATCATCAAGATCTTAAG 300  
DB 241 GACTCAAGAAAGACCTTTGGAGCAAGAAACCTTCCCTGATCATCAAGATCTTAAG 300  
QY 301 ATGAAAGACTCAGATCTTACATCTGTGAAGTGGAGGACCAAGAGAGGTCGAATTG 360  
DB 301 ATGAAAGACTCAGATCTTACATCTGTGAAGTGGAGGACCAAGAGAGGTCGAATTG 360  
QY 361 CTAGTGTGGATTGACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 420  
DB 361 CTAGTGTGGATTGACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 420  
QY 421 CTGACCTTGAAGACCCCTGATGATGAGCCCTCAGTCAATGTAGAGTCCAAAGGGGT 480  
DB 421 CTGACCTTGAAGACCCCTGATGATGAGCCCTCAGTCAATGTAGAGTCCAAAGGGGT 480  
QY 481 AAAAATACAGGGGGGAAAGACCTCTCCGTCTCAGCTGAGCTCCAGATAGTGGC 540  
DB 481 AAAAATACAGGGGGGAAAGACCTCTCCGTCTCAGCTGAGCTCCAGATAGTGGC 540  
QY 541 ACCTGACATGCACTGCTTTCAGAAACAGAAAGTGAATTCAAATAGACATCTGTG 600  
DB 541 ACCTGACATGCACTGCTTTCAGAAACAGAAAGTGAATTCAAATAGACATCTGTG 600  
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
DB 601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626

## RESULT 12

US-09-243-008-1  
Sequence 1, Application US/09243008  
Publication No. US20040005334A1  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian et al.  
TITLE OF INVENTION: Redirection of Cellular Immunity by  
Receptor Chimeras  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
SOFTWARE: WordPerfect (Version 5.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/243,008  
FILING DATE: 02-Feb-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/394,176  
FILING DATE: SEPTEMBER 11,1995  
APPLICATION NUMBER: 08/203,866  
FILING DATE: February 28, 1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: March 7, 1991  
APPLICATION NUMBER: 07/665,961  
ATTORNEY/AGENT INFORMATION:  
NAME: Karen F. Lech, Ph.D  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/270001

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1728 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-243-008-1

Query Match 43.2%; Score 622.8; DB 11; Length 1728;

Best Local Similarity 99.7%; Pred. No. 2,1e-170;

Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGAGAGTCCCTTTTAAAGCACTTCTGTGCTGCAACTGGCGCTCTCCCA 60  
DB 1 ATGAACCGGAGAGTCCCTTTTAAAGCACTTCTGTGCTGCAACTGGCGCTCTCCCA 60  
QY 61 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
DB 61 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
QY 121 TGTACAGCTTCCAGAGAAAGAGATCAATTCACATGAGAAATCTCAACAGATAAG 180  
DB 121 TGTACAGCTTCCAGAGAAAGAGATCAATTCACATGAGAAATCTCAACAGATAAG 180  
QY 181 ATTCTGGGAAATCAGGCTCTCTTTTAACTAAAGTCCATCAAGTGAATGATCGGCT 240  
DB 181 ATTCTGGGAAATCAGGCTCTCTTTTAACTAAAGTCCATCAAGTGAATGATCGGCT 240  
QY 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGCAAGAGAGAGTGCATTAAG 360  
DB 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGCAAGAGAGAGTGCATTAAG 360  
QY 361 CTAGTGTTCGATTTGACTGCACTGCAACCCAGCTTTCAGGGGAGAGCTTACC 420  
DB 361 CTAGTGTTCGATTTGACTGCACTGCAACCCAGCTTTCAGGGGAGAGCTTACC 420  
QY 421 CTGACCTTGGAGAGCCCTCTGTGATGAGCCCTCACTGATGATGAGATCCAGGGGT 480  
DB 421 CTGACCTTGGAGAGCCCTCTGTGATGAGCCCTCACTGATGATGAGATCCAGGGGT 480  
QY 481 AAAAACAATACAGGGGGGAGAGACCTCTCCGTCTCAGCTGAGCTCCAGAGATAGTGC 540  
DB 481 AAAAACAATACAGGGGGGAGAGACCTCTCCGTCTCAGCTGAGCTCCAGAGATAGTGC 540  
QY 541 ACCTGACATGACATGCTCTTTCAGAAACAGAGAGAGTGAATCAAAATAGACATCGTG 600  
DB 541 ACCTGACATGACATGCTCTTTCAGAAACAGAGAGAGTGAATCAAAATAGACATCGTG 600  
QY 601 GTGCTAGCTTTCCAGAAAGGCTTCGG 626  
DB 601 GTGCTAGCTTTCCAGAAAGGCTTCAG 626

## RESULT 13

US-09-891-119A-8

; Sequence 8, Application US/09891119A

; Publication No. US20040013683A1

; GENERAL INFORMATION:

; APPLICANT: Maddon, Paul J.

; TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4

; FILE REFERENCE: 24577-CY-B

; CURRENT APPLICATION NUMBER: US/09/891,119A

; CURRENT FILING DATE: 2001-06-25

## NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 8

; LENGTH: 1742

; TYPE: DNA

; ORGANISM: Human

US-09-891-119A-8

Query Match 43.2%; Score 622.8; DB 11; Length 1742;

Best Local Similarity 99.7%; Pred. No. 2,1e-170;

Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGAGAGTCCCTTTTAAAGCACTTCTGTGCTGCAACTGGCGCTCTCCCA 60  
DB 76 ATGAACCGGAGAGTCCCTTTTAAAGCACTTCTGTGCTGCAACTGGCGCTCTCCCA 135  
QY 61 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
DB 136 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAGGGGATACAGTGAACCTGACC 195  
QY 121 TGTACAGCTTCCAGAGAAAGAGATCAATTCACATGAGAAATCTCAACAGATAAG 180  
DB 196 TGTACAGCTTCCAGAGAAAGAGATCAATTCACATGAGAAATCTCAACAGATAAG 255  
QY 181 ATTCTGGGAAATCAGGCTCTCTTTTAACTAAAGTCCATCAAGTGAATGATCGGCT 240  
DB 256 ATTCTGGGAAATCAGGCTCTCTTTTAACTAAAGTCCATCAAGTGAATGATCGGCT 315  
QY 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 316 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 375  
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGCAAGAGAGAGTGCATTAAG 360  
DB 376 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGCAAGAGAGAGTGCATTAAG 435  
QY 361 CTAGTGTTCGATTTGACTGCACTGCAACCCAGCTTTCAGGGGAGAGCTTACC 420  
DB 436 CTAGTGTTCGATTTGACTGCACTGCAACCCAGCTTTCAGGGGAGAGCTTACC 495  
QY 421 CTGACCTTGGAGAGCCCTCTGTGATGAGCCCTCACTGATGATGAGATCCAGGGGT 480  
DB 496 CTGACCTTGGAGAGCCCTCTGTGATGAGCCCTCACTGATGATGAGATCCAGGGGT 555  
QY 481 AAAAACAATACAGGGGGGAGAGACCTCTCCGTCTCAGCTGAGCTCCAGAGATAGTGC 540  
DB 556 AAAAACAATACAGGGGGGAGAGACCTCTCCGTCTCAGCTGAGCTCCAGAGATAGTGC 615  
QY 541 ACCTGACATGACATGCTCTTTCAGAAACAGAGAGAGTGAATCAAAATAGACATCGTG 600  
DB 616 ACCTGACATGACATGCTCTTTCAGAAACAGAGAGAGTGAATCAAAATAGACATCGTG 675  
QY 601 GTGCTAGCTTTCCAGAAAGGCTTCGG 626  
DB 676 GTGCTAGCTTTCCAGAAAGGCTTCAG 701

## RESULT 14

US-10-024-329-1

; Sequence 1, Application US/10024329

; Publication No. US20030157063A1

; GENERAL INFORMATION:

; APPLICANT: SANHADJI, Kamel

; APPLICANT: TOURAINE, Jean-Louis

; APPLICANT: LEROY, Pierre

; APPLICANT: MENTALI, Maïd

; TITLE OF INVENTION: Gene therapy using anti-gp41 antibody and cd4 immunoadhesin

; FILE REFERENCE: 109993

; CURRENT APPLICATION NUMBER: US/10/024,329

; CURRENT FILING DATE: 2001-12-21

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

LENGTH: 1745  
TYPE: DNA  
ORGANISM: human SCD4  
US-10-024-329-1

Query Match 43.2%; Score 622.8; DB 16; Length 1745;  
Best Local Similarity 99.7%; Pred. No. 2.1e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGACCGGGAGTCCCTTTTAAAGCACTTCTGTGTGCAACTGGCGCTCTCCCA 60
DB 79 ATGACCGGGAGTCCCTTTTAAAGCACTTCTGTGTGCAACTGGCGCTCTCCCA 138
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAGGGATACATGGAATGACC 120
DB 139 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAGGGATACATGGAATGACC 198
QY 121 TGTACGCTTCCAGAAAGAAAGCATATTCACAGTGAATACTCAACCAAGTAAAG 180
DB 199 TGTACGCTTCCAGAAAGAAAGCATATTCACAGTGAATACTCAACCAAGTAAAG 258
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATAAGTTCATCAAGCTGAATGATCGGCT 240
DB 259 ATTCTGGGAATCAGGGCTCTTCTTAATAAGTTCATCAAGCTGAATGATCGGCT 318
QY 241 GACTCAAGAAAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 319 GACTCAAGAAAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 378
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGGAGGTGCAATTG 360
DB 379 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGGAGGTGCAATTG 438
QY 361 CTAGTGTGGATTTGACTGCAACTGTGACCACTGCTTCAAGGGGCAAGCTGACC 420
DB 439 CTAGTGTGGATTTGACTGCAACTGTGACCACTGCTTCAAGGGGCAAGCTGACC 498
QY 421 CTGACCTTGAAGACCCCTGCTGATGAGCCCTCACTGCAATGTAGGATCCAGGGGT 480
DB 499 CTGACCTTGAAGACCCCTGCTGATGAGCCCTCACTGCAATGTAGGATCCAGGGGT 558
QY 481 AAAAACAATCAGGGGGGGAAGACCTCTCGGTGTCAAGTGGAGCTCCAGATAGTGGC 540
DB 559 AAAAACAATCAGGGGGGGAAGACCTCTCGGTGTCAAGTGGAGCTCCAGATAGTGGC 618
QY 541 ACCTGACATGCACTGTCTTGCAAAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 600
DB 619 ACCTGACATGCACTGTCTTGCAAAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 678
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
DB 679 GTGCTAGCTTTCCAGAAAGGCTCCAG 704
```

RESULT 15  
US-10-207-655-169

Sequence 169, Application US/10207655  
Publication No. US20030118592A1  
GENERAL INFORMATION:  
APPLICANT: Ledbetter, Jeffrey A.  
APPLICANT: Hayden-Ledbetter, Martha S.  
TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS  
FILE REFERENCE: 390069.401C1  
CURRENT APPLICATION NUMBER: US/10/207,655  
CURRENT FILING DATE: 2002-07-25  
NUMBER OF SEQ ID NOS: 426  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 169  
LENGTH: 3084  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-207-655-169

Query Match 43.2%; Score 622.8; DB 15; Length 3084;  
Best Local Similarity 99.7%; Pred. No. 2.5e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGACCGGGAGTCCCTTTTAAAGCACTTCTGTGTGCAACTGGCGCTCTCCCA 60
DB 153 ATGACCGGGAGTCCCTTTTAAAGCACTTCTGTGTGCAACTGGCGCTCTCCCA 212
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAGGGATACATGGAATGACC 120
DB 213 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAGGGATACATGGAATGACC 272
QY 121 TGTACGCTTCCAGAAAGAAAGCATATTCACAGTGAATACTCAACCAAGTAAAG 180
DB 273 TGTACGCTTCCAGAAAGAAAGCATATTCACAGTGAATACTCAACCAAGTAAAG 332
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATAAGTTCATCAAGCTGAATGATCGGCT 240
DB 333 ATTCTGGGAATCAGGGCTCTTCTTAATAAGTTCATCAAGCTGAATGATCGGCT 392
QY 241 GACTCAAGAAAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 393 GACTCAAGAAAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 452
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGGAGGTGCAATTG 360
DB 453 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGGAGGTGCAATTG 512
QY 361 CTAGTGTGGATTTGACTGCAACTGTGACCACTGCTTCAAGGGGCAAGCTGACC 420
DB 513 CTAGTGTGGATTTGACTGCAACTGTGACCACTGCTTCAAGGGGCAAGCTGACC 572
QY 421 CTGACCTTGAAGACCCCTGCTGATGAGCCCTCACTGCAATGTAGGATCCAGGGGT 480
DB 573 CTGACCTTGAAGACCCCTGCTGATGAGCCCTCACTGCAATGTAGGATCCAGGGGT 632
QY 481 AAAAACAATCAGGGGGGGAAGACCTCTCGGTGTCAAGTGGAGCTCCAGATAGTGGC 540
DB 633 AAAAACAATCAGGGGGGGAAGACCTCTCGGTGTCAAGTGGAGCTCCAGATAGTGGC 692
QY 541 ACCTGACATGCACTGTCTTGCAAAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 600
DB 693 ACCTGACATGCACTGTCTTGCAAAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 752
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
DB 753 GTGCTAGCTTTCCAGAAAGGCTCCAG 778
```

Search completed: September 20, 2005, 16:05:03  
Job time : 2461 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2005, 08:29:42 ; Search time 5142 Seconds  
(without alignments)  
10659.772 Million cell updates/sec

Title: US-09-936-702-4

Perfect score: 1440  
Sequence: 1 atgaaccg99gagtcctt.....gtcagcagtaactg9 1440

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	622.8	43.2	783	6	CD609192 56069931J
2	622.8	43.2	932	5	BX438745 BX438745
3	610.2	42.4	816	4	BI838409 603083373
4	599.4	41.6	822	5	BX457159 BX457159
5	598.8	41.6	801	4	BI822118 603039813
6	589	40.9	840	4	BI838893 603087279
7	577.4	40.1	791	4	BI819727 603041415
8	555.4	38.6	557	7	CV029573 8403_F011
9	539	37.4	713	5	BX437619 BX437619
10	519.4	36.1	723	4	BI915265 603184313
11	504	35.0	679	6	CD609191 56069931H
12	504	35.0	732	1	AU141298 AU141298
13	488.8	33.9	754	7	CO246446 AGENCOURT
14	488.4	33.9	710	7	CF125444 UT-HF-BL0
15	483.8	33.6	694	6	CB052686 NISC_g109
16	480.6	33.4	803	7	CO249540 AGENCOURT
17	466	32.4	809	7	CO248894 AGENCOURT
18	457.4	31.8	748	5	BX378811 BX378811
19	452.8	31.4	821	1	CO245985 AGENCOURT
20	437.8	30.4	584	5	BP303735 BP303735
21	435.8	29.8	580	5	BP345966 BP345966
22	429.4	28.9	616	7	CF125140 UT-HF-BL0
23	416.4	28.9	584	5	BP287350 BP287350
24	416.4	28.9	585	5	BP286919 BP286919

25	407.6	28.3	751	7	CO248544 AGENCOURT
26	402.6	28.0	687	2	AW752367 RC2-CT020
27	386.8	26.9	583	5	BP361417 BP361417
28	386.4	26.8	582	5	BP339950 BP339950
29	385.8	26.8	583	5	BP287803 BP287803
30	383.4	26.6	579	5	BP345874 BP345874
31	383.4	26.6	582	5	BP300162 BP300162
32	374.6	26.0	767	1	AU136822 AU136822
33	366.6	25.5	686	6	CB052685 NISC_g109
34	364.2	25.3	755	7	CO246525 AGENCOURT
35	358.4	24.9	564	5	BP277052 BP277052
36	354.8	24.6	638	2	AW752358 RC2-CT020
37	345	24.0	893	1	AL553628 AL553628
38	328.8	22.8	707	7	CK835626 4060108 B
39	325.4	22.6	581	5	BP367248 BP367248
40	324.8	22.6	582	5	BP344125 BP344125
41	302.8	21.0	567	6	CB997693 AGENCOURT
42	299.4	20.8	583	5	BP344191 BP344191
43	261.2	18.1	336	2	AW407282 UT-HF-BL0
44	261.2	18.1	374	2	AW406380 UT-HF-BL0
45	261.2	18.1	733	4	BG685053 BG685053

#### ALIGNMENTS

RESULT 1  
LOCUS CD609192 783 bp mRNA linear EST 12-JAN-2004  
DEFINITION 56069931J1 FLP Homo sapiens cDNA, mRNA sequence.  
ACCESSION CD609192  
VERSION CD609192.1 GI:40257455  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.  
1 (bases 1 to 783)  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE Circular rapid amplification of cDNA ends for high-throughput  
extension cloning of partial genes  
JOURNAL Genomics 84 (1), 205-210 (2004)  
COMMENT Contact: Fu GK  
Incyte Genomics, Inc.  
3160 Porter Dr., Palo Alto, CA 94304, USA  
Tel: 6508454102  
Email: gfu@incyte.com.

FEATURES  
source Location/Qualifiers  
1..783  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="FLP"  
/note="Vector: pDrive Cloning Vector"

#### ORIGIN

Query Match 43.2%; Score 622.8; DB 6; Length 783;  
Best Local Similarity 99.7%; Pred. No. 2.5e-152;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGACCGGAGTCCCTTTAGGCACTTGCTTGCTGCTGCACTGGCGCTCTCCCA	60
DB	726	ATGAACCGGAGATCCCTTTAGGCACTTGCTTGCTGCTGCACTGGCGCTCTCCCA	667
QY	61	GCAAGCACTGAGGAAAGAGTGTGCTGGGCAAAAAGGGATACAGTGAACCTGACC	120
DB	666	GCAAGCACTGAGGAAAGAGTGTGCTGGGCAAAAAGGGATACAGTGAACCTGACC	607
QY	121	TGACAGCTTCCAGAAAGAGATCAATTCATGGAATAATCCCAACGATTAAG	180
DB	606	TGACAGCTTCCAGAAAGAGATCAATTCATGGAATAATCCCAACGATTAAG	547
QY	181	ATTCTGGAAATCAGGCTCTTTAAAGTTCATCAAGCTGAATGATCGCGCT	240

```

|||||
Db 546 ATCTGGGAATCGGGCTCTCTTCTTAACATAAGGCTCATCCAGCTGAATGATCGGCT 487
|||
Qy 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCGTCATCATCAAGATCTTAAG 300
|||
Db 486 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCGTCATCATCAAGATCTTAAG 427
|||
Qy 301 ATGAAAGACTCAGATTAATTAATCTGTGAAAGTGAGGACCAAGAAAGAGAGTGCATTTG 360
|||
Db 426 ATGAAAGACTCAGATTAATTAATCTGTGAAAGTGAGGACCAAGAAAGAGAGTGCATTTG 367
|||
Qy 361 CTGAGTTCGGATTGACTGCAACTCTGACACCCACCTGCTTCCGGGGCAGAGCTTACC 420
|||
Db 366 CTGAGTTCGGATTGACTGCAACTCTGACACCCACCTGCTTCCGGGGCAGAGCTTACC 307
|||
Qy 421 CTGACCTTGGAGAGCCCGCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGGT 480
|||
Db 306 CTGACCTTGGAGAGCCCGCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGGT 247
|||
Qy 481 AAAAACAATACAGGGGGGAAAGACCTCTCGGTGTCTGAGCTGAGCTCCAGATATGTC 540
|||
Db 246 AAAAACAATACAGGGGGGAAAGACCTCTCGGTGTCTGAGCTGAGCTCCAGATATGTC 187
|||
Qy 541 ACCTGGACATGCACTGTCTTGAGAGAACCAAGAAAGTGAGTCAAAATAGACATGTC 600
|||
Db 186 ACCTGGACATGCACTGTCTTGAGAGAACCAAGAAAGTGAGTCAAAATAGACATGTC 127
|||
Qy 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
|||
Db 126 GTGCTAGCTTTCCAGAAAGGCTCCAG 101
|||

RESULT 2
LOCUS BX438745 932 bp mRNA linear EST 05-MAY-2004
DEFINITION BX438745 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE002Y001
ACCESSION BX438745
VERSION BX438745.2 GI:47035651
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 932)
L1,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30787731.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Creteilux, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 6485.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna7s=CS0DE002AD01Q1&c=6485.r.

FEATURES
source
1..932
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE002Y001"
/tissue_type="PLACENTA"
/clone_id="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector."

```

```

ORIGIN Library was not normalized."
Query Match 43.2%; Score 622.8; DB 5; Length 932;
Best Local Similarity 99.7%; Pred. No. 2,6e-152;
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAAACGGGAGATCCCTTTTAGGCACTTCTGTGCTGCACTGCGCTCTCCCA 60
|||
Db 180 ATGAAACGGGAGATCCCTTTTAGGCACTTCTGTGCTGCACTGCGCTCTCCCA 239
|||
Qy 61 GCAGCCACTCAGGGAAGAAAGTGCTGGGCAAAAAGGGGATCACTGGAACCTGACC 120
|||
Db 240 GCAGCCACTCAGGGAAGAAAGTGCTGGGCAAAAAGGGGATCACTGGAACCTGACC 299
|||
Qy 121 TGTACAGCTTCCAGAAAGAGATACATTCATCTGCAAAAACTCCAACAGATAAG 180
|||
Db 300 TGTACAGCTTCCAGAAAGAGATACATTCATCTGCAAAAACTCCAACAGATAAG 359
|||
Qy 181 ATTCGGGAATACAGGGCTCTCTTTAACTAAAGTCAATCCAAAGTGAATGCGCT 240
|||
Db 360 ATTCGGGAATACAGGGCTCTCTTTAACTAAAGTCAATCCAAAGTGAATGCGCT 419
|||
Qy 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCGTCATCAAGATCTTAAG 300
|||
Db 420 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCGTCATCAAGATCTTAAG 479
|||
Qy 301 ATGAAAGACTCAGATTAATTAATCTGTGAAAGTGAGGACCAAGAAAGAGGTGCAATTG 360
|||
Db 480 ATGAAAGACTCAGATTAATTAATCTGTGAAAGTGAGGACCAAGAAAGAGGTGCAATTG 539
|||
Qy 361 CTGAGTTCGGATTGACTGCAACTCTGCAACCCACCTGCTTCCAGGGCAGAGCTTACC 420
|||
Db 540 CTGAGTTCGGATTGACTGCAACTCTGCAACCCACCTGCTTCCAGGGCAGAGCTTACC 599
|||
Qy 421 CTGACCTTGGAGAGCCCGCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGGT 480
|||
Db 600 CTGACCTTGGAGAGCCCGCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGGT 659
|||
Qy 481 AAAAACAATACAGGGGGGAAAGACCTCTCGGTGTCTGAGCTGAGCTCCAGATATGTC 540
|||
Db 660 AAAAACAATACAGGGGGGAAAGACCTCTCGGTGTCTGAGCTGAGCTCCAGATATGTC 719
|||
Qy 541 ACCTGGACATGCACTGTCTTGAGAGAACCAAGAAAGTGAGTCAAAATAGACATGTC 600
|||
Db 720 ACCTGGACATGCACTGTCTTGAGAGAACCAAGAAAGTGAGTCAAAATAGACATGTC 779
|||
Qy 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
|||
Db 780 GTGCTAGCTTTCCAGAAAGGCTCCAG 805
|||

RESULT 3
LOCUS BI838409 816 bp mRNA linear EST 04-OCT-2001
DEFINITION 603083373F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:522323 5',
ACCESSION BI838409
VERSION BI838409.1 GI:15949959
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 816)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```



DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
<http://image.jnl.gov>  
 Plate: LAM11559 row: b column: 20  
 High quality sequence stop: 803.  
 Location/Qualifiers

FEATURES  
 source  
 1. 816

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5222323"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH MGC 120"  
 /note="Organ: pooled pancreas and spleen; Vector:  
 pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
 source anonymous pool of spleen and pancreas from 28 yo  
 male. Library is oligo-dT primed and directionally cloned  
 (EcoRV site is destroyed upon cloning). Average insert  
 size 1.5 kb, insert size range 1-2.5 kb. Library is  
 normalized and enriched for full-length clones and was  
 constructed by C. Gruber (Invitrogen). Research Genetics  
 tracking code 025. Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 42.4%; Score 610.2; DB 4; Length 816;

Best Local Similarity 99.4%; Pred. No. 5.2e-149; Indels 1; Gaps 1;

Matches 622; Conservative 0; Mismatches 3;

QY 1 ATGAACCGGGGAGTCCCTTTAGGCACTTGTGTGCTGCAACTGGCGCTCTCCCA 60  
 DB 167 ATGAACCGGGGAGTCCCTTTAGGCACTTGTGTGCTGCAACTGGCGCTCTCCCA 226

QY 61 GCAGCCCTCAGGAAAGAAAGTGTCTGGGCAAAAAGGGATACAGTGAATCACC 120

DB 227 GCAGCCCTCAGGAAAGAAAGTGTCTGGGCAAAAAGGGATACAGTGAATCACC 286

QY 121 TGTACAGCTTCCGAGAAAGAGCATCAATTCACATGAAATCAACAGATAAG 180

DB 287 TGTACAGCTTCCGAGAAAGAGCATCAATTCACATGAAATCAACAGATAAG 346

QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATAAGGTCATCAAGCTGAATGATCGGCT 240

DB 347 ATTCTGGGAATCAGGGCTCTTCTTAATAAGGTCATCAAGCTGAATGATCGGCT 406

QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300

DB 407 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 466

QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGACCAAGAGAGAGGTGCAATTG 360

DB 467 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGACCAAGAGAGAGGTGCAATTG 526

QY 361 CTAGTGTTCGATTGACTGCACTGACACCCACCTGCTTCAGGGGACAGGCTGACC 420

DB 527 CTAGTGTTCGATTGACTGCACTGACACCCACCTGCTTCAGGGGACAGGCTGACC 586

QY 421 CTGACCTTGGAGAGCCCCCTGGTGTAGTACCCCTCACTGCAATTTAGAGTCCAGGGGT 480

DB 587 CTGACCTTGGAGAGCCCCCTGGTGTAGTACCCCTCACTGCAATTTAGAGTCCAGGGGT 646

QY 481 AAAAATCATCAGGGGGGAGAACCTCTCC-GTGTCTCAGCTGAGGCTCCAGATTAATGG 539

DB 647 AAAAATCATCAGGGGGGAGAACCTCTCCGATCTTCAGCTGAGGCTCCAGATTAATGG 706

QY 540 CACCTGACATGCACTGTCTTGCAAGAACAGAAAGAGTGAATTTAAATGACATCGT 599

DB 707 CACCTGACATGCACTGTCTTGCAAGAACAGAAAGAGTGAATTTAAATGACATCGT 766

QY 600 GGTGTAGCTTTCCAGAAAGGCTCTCGG 626

DB 767 GGTGTAGCTTTCCAGAAAGGCTCTCGG 793

RESULT 4  
 BX457159 822 bp mRNA linear EST 05-MAY-2004  
 LOCUS BX457159 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP006YP07  
 DEFINITION 5-PRIME, mRNA sequence.

ACCESSION BX457159  
 VERSION BX457159.2 GI:47051693  
 KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 822)  
 Li, W.B., Gruber, C., Jessee, J. and Polyes, D.

Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On May 22, 2003 this sequence version replaced gi:3102083.

COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.

This sequence belongs to sequence cluster 6485.r

For more information about this cluster, see  
<http://www.genoscope.cns.fr/cdna?b=CS0CAP006Y07&id=6485.r>.

## FEATURES

source

1. 822

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0CAP006YP07"

/issue\_type="THYMUS"

/clone\_lib="Homo sapiens THYMUS"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo(dT) primer. Five prime end enriched,  
 double-strand cDNA was digested with Not I and cloned into  
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
 Library was not normalized."

## ORIGIN

Query Match 41.6%; Score 599.4; DB 5; Length 822;

Best Local Similarity 99.2%; Pred. No. 3.6e-146; Indels 2; Gaps 2;

Matches 622; Conservative 1; Mismatches 2;

QY 1 ATGAACCGGGGAGTCCCTTTAGGCACTTGTGTGCTGCAACTGGCGCTCTCCCA 60

DB 183 ATGAACCGGGGAGTCCCTTTAGGCACTTGTGTGCTGCAACTGGCGCTCTCCCA 242

QY 61 G-CAGCCACTCAGGAAAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAATCAG 119

DB 243 G-CAGCCACTCAGGAAAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAATCAG 302

QY 120 CTGTACAGCTTCCGAGAAAGAGCATCAATTCACCTGGAATAATCCAACTGAATTA 179

DB 303 CTGTACAGCTTCCGAGAAAGAGCATCAATTCACCTGGAATAATCCAACTGAATTA 362

QY 180 GATTCTGGGAAATCAGGGGCTCTTCTTAATCTAAAGGTCATCCAGCTGAATGATGCGC 239

DB 363 GATTCTGGGAAATCAGGGGCTCTTCTTAATCTAAAGGTCATCCAGCTGAATGATGCGC 422

QY 240 TGACTCAAGAAAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAA 299

DB 423 TGACTCAAGAAAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAA 482

QY 300 GATAGAAGACTCAGATTAATCACTGTGTGAAGTGGAGAACCAAGAGAGAGGTGCAATT 359

DB 483 GATAGAAGACTCAGATTAATCACTGTGTGAAGTGGAGAACCAAGAGAGAGGTGCAATT 542

QY 360 GCTAGTGTGGATTGACTGCCAACTCTGAACCCACCTGCTTCAGGGGCGAAGCTTGAC 419  
 DB 543 GCTAGTGTGGATTGACTGCCAACTCTGAACCCACCTGCTTCAGGGGCGAAGCTTGAC 602  
 QY 420 CCTGACCTTGGAGAGCCCCCTGGTAGTACCCTCGATGCAATGTAGAGTCCAAAGGGG 479  
 DB 603 CCTGACCTTGGAGAGCCCCCTGGTAGTACCCTCGATGCAATGTAGAGTCCAAAGGGG 662  
 QY 480 TAAAAACATACAGAGGGGGGAGAGCCCTCTCGTGTCTCACTGAGAGCTCCAGATAGTGG 539  
 DB 663 TAAAAACATACAGAGGGGGGAGAG-CTCTCGTGTCTCACTGAGAGCTCCAGATAGTGG 721  
 QY 540 CACCTGGACATGACATGTCTTTCAGAGAACCAAGAGAGTGTCAAAATAGACATCGT 599  
 DB 722 CACCTGGACATGACATGTCTTTCAGAGAACCAAGAGAGTGTCAAAATAGACATCGT 781  
 QY 600 GGTGCTAGCTTTCCAGAGAGGCTCCGG 626  
 DB 782 GGTGCTAGCTTTCCAGAGAGGCTCCAG 808

RESULT 5  
 BI822118 801 bp mRNA linear EST 04-OCT-2001  
 LOCUS 603039813F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5180642 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI822118  
 VERSION BI822118.1 GI:15933668  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 801)  
 NIH-MGC http://mgs.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM11450 row: j column: 03  
 High quality sequence stop: 793.  
 Location/Qualifiers

FEATURES  
 source  
 1..801  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5180642"  
 /lab\_host="DH10B"  
 /clone\_id="NIH\_MGC\_115"  
 /note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: ScaRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (ScaRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH\_MGC Library."

ORIGIN  
 Query Match 41.6%; Score 598.8; DB 4; Length 801;  
 Best Local Similarity 99.4%; Pred. No. 5,1e-146;  
 Matches 622; Conservative 0; Mismatches 2; Indels 2; Gaps 2;  
 QY 1 ATGAAACGGGAGTCCCTTTAGGCACTTCTGTGCTGCAACTGGCGCTCTCCCA 60

DB 175 ATGAAACGGGAGTCCCTTTAGGCACTTGTGTGCTCAACTGGCGCTCTCCCA 234  
 QY 61 GCAGCCACTCAGGAGAAAGATGCTGGGCAAAAAAGGGATACAGTGAACAGACC 120  
 DB 235 GCAGCCACTCAGGAGAAAGATGCTGGGCAAAAAAGGGATACAGTGAACAGACC 294  
 QY 121 TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGAAAAAACTCCAAACAGATTAAG 180  
 DB 295 TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGAAAAAACTCCAAACAGATTAAG 354  
 QY 181 ATTCTGGAAATCAGGGCTCTCTTAATAAAGGCTCAAGCTGAATATGCGGCT 240  
 DB 355 ATTCTGGAAATCAGGGCTCTCTTAATAAAGGCTCAAGCTGAATATGCGGCT 414  
 QY 241 GACTCAAGAAAGACCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
 DB 415 GACTCAAGAAAGACCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 474  
 QY 301 ATGAAAGACTCAGATACCTTACATCTGTGAAGTGAAGAGCAGAAAGAGAGTGCATTG 360  
 DB 475 ATGAAAGACTCAGATACCTTACATCTGTGAAGTGAAGAGCAGAAAGAGAGTGCATTG 534  
 QY 361 CTAGTGTGGAGATTGACTGCCAATCTGACACCCACCTGCTTCAGGGGCGAAGCTTGACC 420  
 DB 535 CTAGTGTGGAGATTGACTGCCAATCTGACACCCACCTGCTTCAGGGGCGAAGCTTGACC 594  
 QY 421 CTGACCTTGGAGAGCCCCCTGGTAGTACCCTCGATGCAATGTAGAGTCCAAAGGGGT 480  
 DB 595 CTGACCTTGGAGAGCCCCCTGGTAGTACCCTCGATGCAATGTAGAGTCCAAAGGGGT 654  
 QY 481 AAAAAATACA-GGGGGGAGAACCTCTCTCGTGTCTCAGCTGAGCTCCAGATATGTCG 539  
 DB 655 AAAAAATACAAGGGGGGAGAACCTCTCTCGTGTCTCAGCTGAGCTCCAGATATGTCG 714  
 QY 540 CACCTGGACATGACATGTCTTTCAGAGAACCAAGAGAGTGTCAAAATAGACATCGT 599  
 DB 715 CACCTGGACATGACATGTCTTTCAGAGAACCAAGAGAGTGTCAAAATAGACATCGT 774  
 QY 600 -GGTGCTAGCTTTCCAGAGAGGCTCC 624  
 DB 775 GGTGCTAGCTTTCCAGAGAGGCTCC 800

RESULT 6  
 BI838893 840 bp mRNA linear EST 04-OCT-2001  
 LOCUS 603087279F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5226427 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI838893  
 VERSION BI838893.1 GI:15950443  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 840)  
 NIH-MGC http://mgs.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM11569 row: m column: 20  
 High quality sequence start: 11  
 High quality sequence stop: 832.  
 Location/Qualifiers

FEATURES

source

1. .840  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5226427"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_120"  
 /note="Organ: pooled pancreas and spleen; Vector: pcwv-sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 40.9%; Score 589; DB 4; Length 840;  
 Best Local Similarity 98.1%; Pred. No. 2e-143;  
 Matches 617; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 1 ATGAACCGGGAGATCCCTTTTAAAGCACTTGTGCTGCAACTGGGCTCTCCCA 60  
 DB 168 ATGAACCGGGAGATCCCTTTTAAAGCACTTGTGCTGCAACTGGGCTCTCCCA 227  
 QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACATGAACTGACC 120  
 DB 228 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACATGAACTGACC 287  
 QY 121 TGTACAGCTTCCGAAAGAAAGATCAATTCCTGGAATACTCAACAGATTAAG 180  
 DB 288 TGTACAGCTTCCGAAAGAAAGATCAATTCCTGGAATACTCAACAGATTAAG 347  
 QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACTGAAGTCAATCCAGCTGAATGATCGGCT 240  
 DB 348 ATTCTGGGAATCAGGGCTCTTCTTAACTGAAGTCAATCCAGCTGAATGATCGGCT 407  
 QY 241 GACTCAAGAAGAGCTTTGGGACCAAGAACTCCCGCTGATCAATCAAGAACTTAAG 300  
 DB 408 GACTCAAGAAGAGCTTTGGGACCAAGAACTCCCGCTGATCAATCAAGAACTTAAG 467  
 QY 301 ATGAAGACTCAGATCTTACATCTGAAAGTGAAGAGCAAGAGAGAGTCAATTG 360  
 DB 468 ATGAAGACTCAGATCTTACATCTGAAAGTGAAGAGCAAGAGAGAGTCAATTG 527  
 QY 361 CTAGTGTTCGATTAAGTCACTGCAACCACTGCTTCAAGGGCAGAGCTGACC 420  
 DB 528 CTAGTGTTCGATTAAGTCACTGCAACCACTGCTTCAAGGGCAGAGCTGACC 587  
 QY 421 CTGACCTTGAAGAGCCCGCTGAGTGAAGCCCTCAGTGAATGTAGAGTCCAAAGGGGT 480  
 DB 588 CTGACCTTGAAGAGCCCGCTGAGTGAAGCCCTCAGTGAATGTAGAGTCCAAAGGGGT 647  
 QY 481 AAA-AAACATCAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATTAAG 539  
 DB 648 AAAACATCAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATTAAG 707  
 QY 540 CACCTGAGATGCACTGTCTTGAAGAACAGAGAGAGTGAAGTCAAAATGACATGCT 599  
 DB 708 CACCTGAGATGCACTGTCTTGAAGAACAGAGAGAGTGAAGTCAAAATGACATGCT 767  
 QY 600 GG-TGCTAGCTTCCAGAGGCTCCCGGA 627  
 DB 768 GGAGGCTAGCTTCCAGAGGCTCCCGGA 796

RESULT 7  
 BI819727 791 bp mRNA linear EST 04-OCT-2001  
 LOCUS 603041415F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5182274 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI819727

VERSION BI819727.1 GI:15931277  
 EST.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.  
 1 (bases 1 to 791)  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapds-remail.nih.gov](mailto:cgapds-remail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>  
 Plate: LIML1454 row: n column: 03  
 High quality sequence stop: 782.  
 Location/Qualifiers

FEATURES  
 source  
 1. .791  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5182274"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_115"  
 /note="Organ: pooled brain, lung, testis; Vector: pcwv-sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 40.1%; Score 577.4; DB 4; Length 791;  
 Best Local Similarity 99.3%; Pred. No. 2.2e-140;  
 Matches 611; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

QY 1 ATGAACCGGGAGATCCCTTTTAAAGCACTTGTGCTGCAACTGGGCTCTCCCA 60  
 DB 176 ATGAACCGGGAGATCCCTTTTAAAGCACTTGTGCTGCAACTGGGCTCTCCCA 235  
 QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGATACATGAACTGACC 120  
 DB 236 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGATACATGAACTGACC 295  
 QY 121 TGTACAGCTTCCGAAAGAAAGATCAATTCCTGGAATACTCAACAGATTA-A 179  
 DB 296 TGTACAGCTTCCGAAAGAAAGATCAATTCCTGGAATACTCAACAGATTA-A 355  
 QY 180 GATTCTGGGAATCAGGGCTCTTCTTAACTAAGTCCATCCAGCTGAATGATGGCG 239  
 DB 356 GATTCTGGGAATCAGGGCTCTTCTTAACTAAGTCCATCCAGCTGAATGATGGCG 415  
 QY 240 TGACTCAAGAAGAGCTTTGGGACCAAGAACTCCCGCTGATCATCAAGAACTTAA 299  
 DB 416 TGACTCAAGAAGAGCTTTGGGACCAAGAACTTTCCCTGATCATCAAGAACTTAA 475  
 QY 300 GATGAAGACTCAGATTAATTAATCTGTGAAGTGAAGACAGAGAGAGGTGCAATT 359  
 DB 476 GATGAAGACTCAGATTAATTAATCTGTGAAGTGAAGACAGAGAGAGGTGCAATT 535  
 QY 360 GCTAGTGTTCGATTAAGTCACTGCAACCACTGCTTCAAGGGCAGAGCTTGAC 419  
 DB 536 GCTAGTGTTCGATTAAGTCACTGCAACCACTGCTTCAAGGGCAGAGCTTGAC 595

QY 420 CCGTACCTTGG-AGAGCCCCCTGGTAGTACCCCTCAGTGCATGTAGAGATCCAGGG 478  
 Db 596 CCGTACCTTGGCAGAGCCCCCTGGTAGTACCCCTCAGTGCATGTAGAGATCCAGGG 655  
 QY 479 GTAAAAACATACAGGGGGGGAAGACCTCTCCGGTGTACGTGGAGCTCAGAGATAGTG 538  
 Db 656 GTAAAAACATACAGGGGGGGAAGACCTCTCCGGTGTACGTGGAGCTCAGAGATAGTG 715  
 QY 539 GCACCTGAGCA-TCGACTGTCTTGCAGAACACAGAAAGGTGAGTTCAAAATAGACATC 597  
 Db 716 GCACCTGAGCATTCGACTGTCTTGCAGAACACAGAAAGGTGAGTTCAAAATAGACATC 775  
 QY 598 GTGGTGTAGCTTTC 612  
 Db 776 GTGGTGTAGCTTTC 790

RESULT 8  
 CV029573  
 LOCUS 557 bp mRNA linear EST 20-AUG-2004  
 DEFINITION 8403 Full Length cDNA from the Mammalian Gene Collection Homo  
 sapiens cDNA 5' similar to BC025782, mRNA sequence.  
 ACCESSION CV029573  
 VERSION CV029573.1 GI:51487751  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 557)  
 Rial,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S.,  
 Dicot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O.,  
 Clingingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T.,  
 Simmons,B., Segretera,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C.,  
 Vandenhaute,J., Chusick,M.E., Alpala,J.S., Hill,D.E. and Vidal,M.  
 Human ORFeome Version 1.1: a Platform for Reverse Proteomics  
 Genome Res. (2004) In press  
 COMMENT  
 Contact: Vidal M  
 Marc Vidal Laboratory  
 Dana Farber Cancer Institute  
 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA  
 Tel: 617 632 5180  
 Fax: 617 632 5739  
 Email: Marc.Vidal@dfci.harvard.edu  
 ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF  
 results from a PCR reaction using an MGC full-length cDNA as  
 template DNA and ORF specific primers  
 PCR Primers  
 FORWARD: ATGACCGGGAGTCCCT  
 BACKWARD: CAATGGGGTACATGCTTTCGAAA  
 Insert Length: 557 Std Error: 21.00  
 Plate: 11064 row: 03 column: B  
 Seq primer: ACTGGCGGTCTTTTACACGTCGTGACTGGAAAC  
 High quality sequence start: 98  
 High quality sequence stop: 556  
 POLYA=No.  
 location/qualifiers  
 1. 557  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="mixed"  
 /clone\_lib="Full Length cDNA from the Mammalian Gene  
 Collection"  
 /note="Vector: mixed; The ORFs were PCR amplified from the  
 MGC (Mammalian Gene Collection) as of April 2004 and  
 cloned by recombinational Gateway cloning into pDONR223  
 Donor Vector. Reference : MGC (Mammalian Gene Collection)  
 Program Team, Generation and Initial Analysis of more than  
 15,000 Full-length Human and Mouse cDNA Sequences. PNAS,  
 2002, 99(126), 16899-16903"

FEATURES  
 source  
 ORIGIN

Query Match 38.6%; Score 555.4; DB 7; Length 557;  
 Best Local Similarity 99.8%; Pred. No. 1.2e-134;  
 Matches 556; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGAGTGTGCAACTGGCGCTCCCA 60  
 Db 1 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGAGTGTGCAACTGGCGCTCCCA 60  
 QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAACAGACC 120  
 Db 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAACAGACC 120  
 QY 121 TGTACAGCTTCCGAGAAAGACATCAATTCCAGTGAATAAATCTCAACCATTAAG 180  
 Db 121 TGTACAGCTTCCGAGAAAGACATCAATTCCAGTGAATAAATCTCAACCATTAAG 180  
 QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATAAGGTCCATCCAGCTGAATAGCGGCT 240  
 Db 181 ATTCTGGGAATCAGGGCTCTTCTTAATAAGGTCCATCCAGCTGAATAGCGGCT 240  
 QY 241 GACTCAAGAAAGAGCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 300  
 Db 241 GACTCAAGAAAGAGCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 300  
 QY 301 ATGAAGACTCAGATATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 360  
 Db 301 ATGAAGACTCAGATATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 360  
 QY 361 CTAGTGTGGATGCTGCTGCACTGACCACTGCTCAGAGGAGAGCTGAC 420  
 Db 361 CTAGTGTGGATGCTGCTGCACTGACCACTGCTCAGAGGAGAGCTGAC 420  
 QY 421 CTGACCTTGAAGAGCCCCCTGGTAGTACCCCTCAGTGCATGTAGAGTCAAGGGT 480  
 Db 421 CTGACCTTGAAGAGCCCCCTGGTAGTACCCCTCAGTGCATGTAGAGTCAAGGGT 480  
 QY 481 AAAAAATACAGGGGGGGAAGACCTCTCTCGTGTCTCAGCTGAGAGTCCAGATAGTGGC 540  
 Db 481 AAAAAATACAGGGGGGGAAGACCTCTCTCGTGTCTCAGCTGAGAGTCCAGATAGTGGC 540  
 QY 541 ACCTGAGCATGCACTGT 557  
 Db 541 ACCTGAGCATGCACTGT 557

RESULT 9  
 BX437619 713 bp mRNA linear EST 04-MAY-2004  
 LOCUS BX437619  
 DEFINITION BX437619 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP007YD02  
 5-PRIME, mRNA sequence.  
 ACCESSION BX437619  
 VERSION BX437619.2 GI:4699918  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 713)  
 Li,W.B., Gruber,C., Jesse,J. and Polayee,D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On May 15, 2003 this sequence version replaced gi:30771588.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqefgenoscope.cns.fr Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.  
 This sequence belongs to sequence cluster 6485.r  
 For more information about this cluster, see

http://www.genoscope.cns.fr/cdnas=CS0CAP007DB010P1ec=6485.r.  
Location/Qualifiers

## FEATURES

source

1. 713  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0CAP007YD02"  
/issue\_type="THYMUS"  
/clone\_lib="Homo sapiens THYMUS"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

## ORIGIN

Query Match 37.4%; Score 539; DB 5; Length 713;  
Best Local Similarity 99.5%; Pred. No. 2.6e-130;  
Matches 550; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

```

QY 1 ATGAACGGGGAGTCCCTTTAGGCACTTCTGTGGCTGCAACTGGGCTCTCCCA 60
Db 161 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGGCTGCAACTGGGCTCTCCCA 220
QY 61 GCAGCCACTCAGAGGAAAGAGTGTGTGGGCAAAAAAGGGGATAGAGTGAATGACC 120
Db 221 GCAGCCACTCAGAGGAAAGAGTGTGTGGGCAAAAAAGGGGATAGAGTGAATGACC 280
QY 121 TGTACAGCTTCCGAAAGAGATACATTCACCTGAGAAAACTCCAAACAGATTAAG 180
Db 281 TGTACAGCTTCCGAAAGAGATACATTCACCTGAGAAAACTCCAAACAGATTAAG 340
QY 181 ATTCTGGGAATCAGGGCTCTCTTAACTTAAGTCCATCCAGCTGAATGAGGGCT 240
Db 341 ATTCTGGGAATCAGGGCTCTCTTAACTTAAGTCCATCCAGCTGAATGAGGGCT 400
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
Db 401 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 460
QY 301 ATGAAGACTCAGATTAATCTTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 360
Db 461 ATGAAGACTCAGATTAATCTTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 520
QY 361 CTAGTGTTCGATGATCTGCAACTGACCACTGCTTCAAGGGGCAAGGCTTACC 420
Db 521 CTAGTGTTCGATGATCTGCAACTGACCACTGCTTCAAGGGGCAAGGCTTACC 580
QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGAATGAGAGTCAAGGGGT 480
Db 581 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGAATGAGAGTCAAGGGGT 640
QY 481 AAAAACAATACA-GGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATTAAG 539
Db 641 AAAAACAATACAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATTAAG 700
QY 540 CACCTGACATGC 552
Db 701 CACCTGACATGC 713

```

## RESULT 10

LOCUS

BI915265 723 bp mRNA linear EST 16-OCT-2001  
60318431F1 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:524826 5'

DEFINITION

BI915265  
BI915265

ACCESSION

BI915265.1 GI:16179363

VERSION

EST. Homo sapiens (human)

KEYWORDS

ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 723)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@rs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:  
http://image.lnl.gov  
Plate: LLAM1626 row: j column: 03  
High quality sequence stop: 717.

## FEATURES

source

1. 723  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:524826"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_121"  
/note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 36.1%; Score 519.4; DB 4; Length 723;  
Best Local Similarity 99.1%; Pred. No. 3.8e-125;  
Matches 564; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

```

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGGCTGCAACTGGGCTCTCCCA 60
Db 157 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGGCTGCAACTGGGCTCTCCCA 216
QY 61 GCAGCCACTCAGAGG-AAAAGAGTGTGTGGGCAAAAAAGGGGATAGAGTGAATGACC 119
Db 217 GCAGCCACTCAGAGGATAGAGTGTGTGGGCAAAAAAGGGGATAGAGTGAATGACC 276
QY 120 CTGACAGCTTCCGAAAGAGATACATTCACCTGAGAAAACTCCAAACAGATTAAG 179
Db 277 CTGACAGCTTCCGAAAGAGATACATTCACCTGAGAAAACTCCAAACAGATTAAG 336
QY 180 GATTCTGGGAATCAGGGCTCTCTTAACTTAAGTCCATCCAGTGAATGATGCCGC 239
Db 337 GATTCTGGGAATCAGGGCTCTCTTAACTTAAGTCCATCCAGTGAATGATGCCGC 396
QY 240 TGACTCAAGAAAGACCTTTGGGACCAAGGAACCTCCCTGATCATCAAGATCTTAAG 299
Db 397 TGACTCAAGAAAGACCTTTGGGACCAAGGAACCTCCCTGATCATCAAGATCTTAAG 456
QY 300 GATGAAGACTCAGATTAATCTTGTGAAGTGAAGACCAAGAGAGAGTGCATTT 359
Db 457 GATGAAGACTCAGATTAATCTTGTGAAGTGAAGACCAAGAGAGAGTGCATTT 516
QY 360 GCTAGTGTTCGATGATCTGCAACTGACCACTGCTTCAAGGGGCA-GAGCTTGA 418
Db 517 GCTAGTGTTCGATGATCTGCAACTGACCACTGCTTCAAGGGGCAAGGCTTGA 576
QY 419 CCTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGAATGAGAGTCCAAAGGG 478
Db 577 CCTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGAATGAGAGTCCAAAGGG 636
QY 479 GTAAAAACATACAGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATTAAG 538

```

Db 637 GTAAAAACATACA-GGGGGGAAGACCTCTCCGTCTCTCACTGAGCTCCAGATAGTG 635

QY 539 GCACCTGACATGACATGCTCTTGAGAAC 567

Db 696 G-ACCTGACATGACATGCTCTTGAGAAC 723

RESULT 11

LOCUS CD609191 679 bp mRNA linear EST 12-JAN-2004

DEFINITION 5606933.H1 FLP Homo sapiens cDNA, mRNA sequence.

ACCESSION CD609191

VERSION CD609191.1 GI:40257454

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 679) Yang, J., Au-Young, J. and Stuve, L.L. Mammalian Eutheria; Primates; Catarrhini; Homnidae; Homo. Circular rapid amplification of cDNA ends for high-throughput extension cloning of partial genes

Genomics 84 (1), 205-210 (2004)

CONTACT: Fu GK

Incye Genomics, Inc.

3160 Porter Dr., Palo Alto, CA 94304, USA

Tel: 6508454102

Email: gfu@incye.com.

Location/Qualifiers

1..679

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone\_lib="FLP"

/note="Vector: pDrive Cloning Vector"

ORIGIN

Query Match 35.2%; Score 506.4; DB 6; Length 679;

Best Local Similarity 99.8%; Pred. No. 9,7e-122;

Matches 507; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACGGGGAGTCCCTTTTAAAGCACTTCTGTGTGCAACGGGGCTCCCTCCA 60

Db 172 ATGAACGGGGAGTCCCTTTTAAAGCACTTCTGTGTGCAACGGGGCTCCCTCCA 231

QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAATGACC 120

Db 232 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAATGACC 231

QY 121 TGTACAGCTTCCAGAAAGAGATCAATTCCTGAAATTCCTCAACAGATTAAG 180

Db 292 TGTACAGCTTCCAGAAAGAGATCAATTCCTGAAATTCCTCAACAGATTAAG 351

QY 181 ATTCTGGGAATCAGGGCTCTCTTTAAAGTTCATCAAGCTGAATGATCGGCT 240

Db 352 ATTCTGGGAATCAGGGCTCTCTTTAAAGTTCATCAAGCTGAATGATCGGCT 411

QY 241 GACTCAAGAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300

Db 412 GACTCAAGAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 471

QY 301 ATAGAAGCTCAGATTAATCATCTGTGAAGTGAAGACCAAGAAGAGAGTCAATTG 360

Db 472 ATAGAAGCTCAGATTAATCATCTGTGAAGTGAAGACCAAGAAGAGAGTCAATTG 531

QY 361 CTGATGTTGGATTTGAATGCACTGCAACCACTGCTTCAAGGGGCAAGCTGACC 420

Db 532 CTGATGTTGGATTTGAATGCACTGCAACCACTGCTTCAAGGGGCAAGCTGACC 531

QY 421 CTGACCTTGAGAGCCCTCTGTGATGAGCCCTCACTGATGATGATGATGATGATG 480

Db 592 CTGACCTTGAGAGCCCTCTGTGATGAGCCCTCACTGATGATGATGATGATGATG 651

QY 481 AAAAACATACAGGGGGGAAGACCTCT 508

Db 652 AAAAACATACAGGGGGGAAGACCTCT 679

RESULT 12

LOCUS AU141298 732 bp mRNA linear EST 05-AUG-2002

DEFINITION AU141298 THYR01 Homo sapiens cDNA clone THYR01000359 5', mRNA sequence.

ACCESSION AU141298

VERSION AU141298.1 GI:11002819

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 732) Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saico, K., Kawai, Y., Isogai, T. HRI human cDNA project

Unpublished (2000)

CONTACT: Takao Isogai

Genomics Laboratory

Helix Research Institute

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

Location/Qualifiers

1..732

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="THYR01000359"

/issue\_type="Chryoid gland"

/clone\_lib="THYR01"

/note="Vector: pME18SFL3"

ORIGIN

Query Match 35.0%; Score 504; DB 1; Length 732;

Best Local Similarity 97.2%; Pred. No. 4.2e-121;

Matches 522; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 1 ATGAACGGGGAGTCCCTTTTAAAGCACTTCTGTGTGCAACGGGGCTCCCTCCA 60

Db 196 ATGAACGGGGAGTCCCTTTTAAAGCACTTCTGTGTGCAACGGGGCTCCCTCCA 255

QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAATGACC 120

Db 256 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAATGACC 315

QY 121 TGTACAGCTTCCAGAAAGAGATCAATTCCTGAAATTCCTCAACAGATTAAG 180

Db 316 TGTACAGCTTCCAGAAAGAGATCAATTCCTGAAATTCCTCAACAGATTAAG 375

QY 181 ATTCTGGGAATCAGGGCTCTCTTTAAAGTTCATCAAGCTGAATGATGCGCT 240

Db 376 ATTCTGGGAATCAGGGCTCTCTTTAAAGTTCATCAAGCTGAATGATGCGCT 435

QY 241 GACTCAAGAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300

Db 436 GACTCAAGAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 495

QY 301 ATAGAAGCTCAGATTAATCATCTGTGAAGTGAAGACCAAGAAGAGAGTGAATTG 360

Db 496 ATAGAAGCTCAGATTAATCATCTGTGAAGTGAAGACCAAGAAGAGAGTGAATTG 555

QY 361 CTAGTGTTCGATTGACTGCCAATCTGACACCCACCTGCTTACGGGGAGAGGCTGACC 420  
 DB 556 CTAGTGTTCGATTGACTGCCAATCTGACACCCACCTGCTTACGGGGAGAGGCTGACC 615  
 QY 421 CTGACCTTGGAGAGCCCCCTGCTGATGAGCCCTCACTGATGATAGAGATCCAGGGGT 480  
 DB 616 CTGACCTTGGAGAGCCCCCTGCTGATGAGCCCTCACTGATGATAGAGATCCAGGGGT 675  
 QY 481 AAAAACAATACAGGGGGGGG--AAGACCTCTCCGCTGCTGAGCTGAGACTCCAGGATA 535  
 DB 676 AAAAACAATACAGGGGGGGGAAACCTTTTCGGGGTCTAACTTGAGCTTCANAGATA 732

RESULT 13  
 CO246446 754 bp mRNA linear EST 23-JUN-2004  
 LOCUS AGENCOURT\_26517764 NIH\_MGC\_212 Homo sapiens cDNA clone

ACCESSION CO246446  
 VERSION CO246446  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 754)  
 NIH-MGC http://mgs.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 CONTACT: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: Mary Hendrix  
 CDNA Library Preparation: M. Bento Soares, University of Iowa  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: NDAMI165 row: O column: 18  
 High quality sequence stop: 507.

FEATURES  
 source location/Qualifiers

1..754  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30924065"  
 /issue\_type="Chondrosarcoma Lung Metastasis cell lines"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_id="NIH\_MGC\_212"  
 /note="Organ: Lung; Vector: pYX-Asc; Site: 1: Scori I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fractionation, ligated with Scori I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GATAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN

Query Match 33.9%; Score 488.8; DB 7; Length 754;  
 Best Local Similarity 96.6%; Pred. No. 4.2e-117;  
 Matches 510; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 1 ATGAACGGGAGTCCCTTTAGGCACTTGTGCTGCTGCACTGGGGCTCTCCCA 60  
 DB 173 ATGAACGGGAGTCCCTTTAGGCACTTGTGCTGCTGCACTGGGGCTCTCCCA 232

QY 61 GCAGCCACTCAGGAGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATA CAGTGAACCTGACC 120  
 DB 233 GCAGCCACTCAGGAGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATA CAGTGAACCTGACC 292  
 QY 121 TGTACAGCTTCCCAAGAAAGAGACATACATTCACCTGAAAAAATCTCAACGATTAAG 180  
 DB 293 TGTACAGCTTCCCAAGAAAGAGACATACATTCACCTGAAAAAATCTCAACGATTAAG 352  
 QY 181 ATTCTGGGAAATAGAGGCTCTCTTAACTAAAGATCCATCCAGGCTGAATGCGGCT 240  
 DB 353 ATTCTGGGAAATAGAGGCTCTCTTAACTAAAGATCCATCCAGGCTGAATGCGGCT 412  
 QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 300  
 DB 413 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 472  
 QY 301 ATGAAGACTCAGATACCTTACATCTGTGAAGTGGAGACCAAGAAAGAGGTGCAATTG 360  
 DB 473 ATGAAGACTCAGATACCTTACATCTGTGAAGTGGAGACCAAGAAAGAGGTGCAATTG 532  
 QY 361 CTGAGTGTGGATTGACTGCACTGCACTGCACTGCACTGCTTCAAGGGGAGAGCTGACC 420  
 DB 533 CTGAGTGTGGATTGACTGCACTGCACTGCACTGCACTGCTTCAAGGGGAGAGCTGACC 592  
 QY 421 CTGACCTTGGAGAGCCCCCTGCTGATGAGCCCTCACTGATGATAGAGTCCAGGGGT 480  
 DB 593 CTGACCTTGGAGAGCCCCCTGCTGATGAGCCCTCACTGATGATAGAGTCCAGGGGT 652  
 QY 481 AAAAACAATACAGGGGGGAAAGACCTCTCTGCTGCTGAGTCCAGGCTC 528  
 DB 653 AAAAACAATACAGGGGGGAAAGACCTCTCTGCTGCTGAGTCCAGGCTC 699

RESULT 14  
 CF125444 710 bp mRNA linear EST 05-AUG-2003  
 LOCUS UI-HF-E10-avo-e-23-0-UI.r1 NIH\_MGC\_212 Homo sapiens cDNA clone  
 ACCESSION CF125444  
 VERSION CF125444  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 710)  
 Bernaldo, M.F., Lennon, G. and Soares, M.B.  
 Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8889548

COMMENT Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Mary Hendrix  
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
 http://genome.uiowa.edu/distribution/humanfl.html  
 Seq primer: pYX-5.

FEATURES  
 source location/Qualifiers

1..710  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30563734"  
 /issue\_type="Chondrosarcoma Lung Metastasis cell lines"

/lab\_host="DH10B (T1 phage resistant)"  
 /clone\_11b="NIH MG-212"  
 /note="Organ: Lung; Vector: pYX-Asc; Site 1: EcoR I;  
 Site 2: Not I; The library was constructed according  
 Bernaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction. Ligated  
 with EcoR I adaptor, digested with Not I and then cloned  
 directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is GATTAAGGCCA. Tissue was provided by Mary Hendrix."

## ORIGIN

Query Match 33.9%; Score 488.4; DB 7; Length 710;  
 Best Local Similarity 98.9%; Pred. No. 5.3e-117;  
 Matches 522; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGTGTGTGCAACTGGCGCTCTCCCA 60  
 DB 186 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGTGTGTGCAACTGGCGCTCTCCCA 245  
 QY 61 GCAGCCACTAGAGGAAAGAAAGTGTCTGGGCAAAAAGGGATACAGTGGAACTGACC 120  
 DB 246 GCAGCCACTAGAGGAAAGAAAGTGTCTGGGCAAAAAGGGATACAGTGGAACTGACC 305  
 QY 121 TGAACAGCTTCCAGAAAGAGAGATACATTCACACTGGAAAACTCCACACAGATTAAG 180  
 DB 306 TGAACAGCTTCCAGAAAGAGAGATACATTCACACTGGAAAACTCCACACAGATTAAG 365  
 QY 181 ATTCTGGGAATCAGGGCTCCTTTTAACTAAAGTTCATCAAGCTGAATGATCGGCT 240  
 DB 366 ATTCTGGGAATCAGGGCTCCTTTTAACTAAAGTTCATCAAGCTGAATGATCGGCT 425  
 QY 241 GATCTCAAGAAAGCTTTGGGACCAAGGAACTTCCCTGTATCATCAAGATCTTAAAG 300  
 DB 426 GATCTCAAGAAAGCTTTGGGACCAAGGAACTTCCCTGTATCATCAAGATCTTAAAG 485  
 QY 301 ATGAAGACTCAGATATTACATCTGTGAAGTGAAGAGAGAGAGAGAGATGCAATTG 360  
 DB 486 ATGAAGACTCAGATATTACATCTGTGAAGTGAAGAGAGAGAGAGATGCAATTG 545  
 QY 361 CTAGTGTTCGATTGACTGTGCAACTGTGACACCACTGTCTTCAAGGGGACGCTGACC 420  
 DB 546 CTAGTGTTCGATTGACTGTGCAACTGTGACACCACTGTCTTCAAGGGGACGCTGACC 605  
 QY 421 CTGACCTTGAAGAGCCCTTGTGTAGAGCCCTCACTGCAATGTAGAAATCCAAAGGGGT 480  
 DB 606 CTGACCTTGAAGAG-CCCTGTGTAGTAG-CCCTCACTGCAATGTAGAAATCCAAAGGGGT 663  
 QY 481 AAAAACAATACAGGGGGAGAGACCTCTCGTGTCTCAGCTGAGCTC 528  
 DB 664 AAAAACAATACAGGGGGAGAG-CTCTCTCCGTGTCTCAGCTGAGCTC 710

RESULT 15  
 LOCUS CB052686 694 bp mRNA linear EST 17-JAN-2003  
 DEFINITION NISC\_gli9e06.y1 NCI\_CGAP\_Le12 Homo sapiens cDNA clone IMAGE:3290578  
 5', mRNA sequence.  
 ACCESSION CB052686  
 VERSION CB052686.1 GI:27790973  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 694)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-r@mail.nih.gov  
 CDNA Library Preparation:  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL  
 Sequencing Center (NISC)  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 info@image.llnl.gov  
 Plate: LLAM8058 row: I column: 11  
 Seq primer: M13Rpi reverse primer (AB1).  
 Location/Qualifiers  
 1. 694

## FEATURES

source

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3290578"  
 /tissue\_type="leiomyosarcoma"  
 /lab\_host="DH10B"  
 /clone\_11b="NCI\_CGAP\_Le12"

/note="Organ: soft tissue; Vector: pYX-Asc (Pharmacia)  
 with a modified polylinker; Site 1: Not I; Site 2: Eco RI;  
 1st strand cDNA was primed with a Not I - oligo(dT) primer  
 [5'-ACTGGAAGAAATCGCGCCGCAATCTTTTCTTTTCTTTT-3'],  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pYX-Asc vector. Library  
 went through one round of normalization. Library  
 constructed by Bento Soares and M. Fatima Bernaldo."

## ORIGIN

Query Match 33.6%; Score 483.8; DB 6; Length 694;  
 Best Local Similarity 99.4%; Pred. No. 8.5e-116;  
 Matches 496; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGTGTGTGCAACTGGCGCTCTCCCA 60  
 DB 196 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGTGTGTGCAACTGGCGCTCTCCCA 255  
 QY 61 GCAGCCACTAGAGGAAAGAAAGTGTCTGGGCAAAAAGGGATACAGTGGAACTGACC 119  
 DB 256 GCAGCCACTAGAGGAAAGAAAGTGTCTGGGCAAAAAGGGATACAGTGGAACTGACC 315  
 QY 120 TGAACAGCTTCCAGAAAGAGAGATACATTCACACTGGAAAACTCCACACAGATTA 179  
 DB 316 TGAACAGCTTCCAGAAAGAGAGATACATTCACACTGGAAAACTCCACACAGATTA 375  
 QY 180 GATTCGGGAATCAGGGCTCCTTTTAACTAAAGTTCATCAAGCTGAATGATCGGCT 239  
 DB 376 GATTCGGGAATCAGGGCTCCTTTTAACTAAAGTTCATCAAGCTGAATGATCGGCT 435  
 QY 421 CTGACCTTGAAGAGCCCTTGTGTAGAGCCCTCACTGCAATGTAGAAATCCAAAGGGGT 480  
 DB 606 CTGACCTTGAAGAG-CCCTGTGTAGTAG-CCCTCACTGCAATGTAGAAATCCAAAGGGGT 663  
 QY 481 AAAAACAATACAGGGGGAGAGACCTCTCGTGTCTCAGCTGAGCTC 528  
 DB 664 AAAAACAATACAGGGGGAGAG-CTCTCTCCGTGTCTCAGCTGAGCTC 594

Search completed: September 20, 2005, 12:03:00



Job time : 5149 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2005, 13:01:41 / Search time 161 Seconds  
(without alignments)  
622.180 Million cell updates/sec

Title: US-09-936-702-3\_COPY\_244\_502

Sequence: 1 QVQLTSGAEVKKPKSSSVK.....RFTFGQTRLEIKLVERGSG 259

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2815

Minimum DB seq length: 259  
Maximum DB seq length: 259

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: geneseqp1980s:\*\n2: geneseqp1990s:\*\n3: geneseqp2000s:\*\n4: geneseqp2001s:\*\n5: geneseqp2002s:\*\n6: geneseqp2003as:\*\n7: geneseqp2003bs:\*\n8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	8.1	259	5	ABP45023 Human Bly
2	21	8.1	259	7	ADG95850 Single ch
3	20	7.7	259	5	ABP45252 Human Bly
4	20	7.7	259	5	ABP45264 Human Bly
5	20	7.7	259	5	ABP45236 Human Bly
6	20	7.7	259	5	ABP45032 Human Bly
7	20	7.7	259	5	ABP45259 Human Bly
8	20	7.7	259	5	ABP45486 Human Bly
9	20	7.7	259	7	ADG96086 Human Bly
10	20	7.7	259	7	ADG96091 Human Bly
11	20	7.7	259	7	ADG96313 Single ch
12	20	7.7	259	7	ADG96063 Single ch
13	20	7.7	259	7	ADG95859 Single ch
14	20	7.7	259	7	ADG96079 Single ch
15	18	6.9	259	5	AU72862 P4-3 sing
16	16	6.2	259	3	AA809775 TWY 30K m
17	16	6.2	259	5	ABP45345 Human Bly
18	16	6.2	259	5	ABP44853 Human Bly
19	16	6.2	259	5	ABP45474 Human Bly
20	16	6.2	259	5	ABP45441 Human Bly
21	16	6.2	259	5	ABP45470 Human Bly
22	16	6.2	259	5	ABP45021 Human Bly
23	16	6.2	259	5	ABP45030 Human Bly
24	16	6.2	259	5	ABP45556 Human Bly
25	16	6.2	259	5	ABP45014 Human Bly

26	16	6.2	259	5	ABP45653 Human Bly
27	16	6.2	259	5	ABP45266 Human Bly
28	16	6.2	259	5	ABP45541 Human Bly
29	16	6.2	259	5	ABP44962 Human Bly
30	16	6.2	259	5	ABP45251 Human Bly
31	16	6.2	259	7	ADG32324 Mouse scf
32	16	6.2	259	7	ADG32322 Mouse scf
33	16	6.2	259	7	ADG35857 Single ch
34	16	6.2	259	7	ADG36383 Single ch
35	16	6.2	259	7	ADG36368 Single ch
36	16	6.2	259	7	ADG36480 Single ch
37	16	6.2	259	7	ADG36268 Single ch
38	16	6.2	259	7	ADG35680 Single ch
39	16	6.2	259	7	ADG36301 Single ch
40	16	6.2	259	7	ADG36093 Single ch
41	16	6.2	259	7	ADG36297 Single ch
42	16	6.2	259	7	ADG36172 Single ch
43	16	6.2	259	7	ADG35841 Single ch
44	16	6.2	259	7	ADG36078 Single ch
45	16	6.2	259	7	ADG35789 Single ch

## ALIGNMENTS

RESULT 1  
ABP45023 standard; protein; 259 AA.

19-AUG-2002 (first entry)  
Human Blys binding scfv seq ID 1034.

DE XX Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;  
KW tumour necrosis factor; B cell proliferation; B cell differentiation;  
KW immunosuppressive; immunostimulant; antirheumatic;  
KW antiAids; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;  
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;  
KW common variable immunodeficiency; acquired immunodeficiency syndrome.  
XX Homo sapiens.  
OS  
XX  
PN WO200202641-A1.  
PD 10-JAN-2002.  
XX  
PF 15-JUN-2001; 2001WO-US019110.  
XX 16-JUN-2000; 2000US-0212210P.  
PR 17-OCT-2000; 2000US-0240816P.  
PR 16-MAR-2001; 2001US-0276248P.  
PR 21-MAR-2001; 2001US-0277379P.  
PR 25-MAY-2001; 2001US-0293499P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.  
XX  
XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;  
XX WPL; 2002-114799/15.  
XX  
PT Antibodies against B lymphocyte Stimulating polypeptides, useful for the  
XX diagnosis and treatment of cancers and immune disorders.  
XX  
PS Claim 1; Page 1635-1636; 3148pp; English.  
XX  
CC This invention describes novel antibodies that immunospecifically bind to  
CC B lymphocyte stimulator (Blys) polypeptides. Blys is a member of the  
CC tumour necrosis factor (TNF) super family and induces B cell  
CC proliferation and differentiation. The antibodies of the invention have  
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,

CC antineumatic and antiAIDS activity and can be used in vaccines to  
 CC inhibit the expression and activity of Blys. The antibodies bind to Blys  
 CC and so may be used to detect and quantitate the presence of Blys in  
 CC biological samples and may be used in this way to diagnose disease  
 CC associated with aberrant expression of Blys. They may also be  
 CC administered to treat diseases associated with aberrant Blys expression  
 CC and activity such as cancer, immune, and autoimmune disorders and  
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,  
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and  
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent  
 CC the antibodies and fragments of the antibodies described in the method of  
 CC the invention

XX Sequence 259 AA;

Query Match 8.1%; Score 21; DB 5; Length 259;

Best Local Similarity 100.0%; Pred. No. 3.7e-10; Mismatches 0; Indels 0; Gaps 0;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 ESGAEVKKPGSSVKSCKASG 26  
 Db 6 ESGAEVKKPGSSVKSCKASG 26

RESULT 2  
 ADG95850  
 ID ADG95850 standard; protein; 259 AA.  
 XX  
 AC ADG95850;

DT 11-MAR-2004 (first entry)

XX Single chain antibody that immunospecifically binds Blys SeqID 1034.

XX  
 DE  
 XX  
 KW antibody; B lymphocyte stimulator; Blys; tumor necrosis factor;  
 KW B cell proliferation; differentiation; scfv; myasthenia gravis;  
 KW multiple sclerosis; aschma; rheumatoid arthritis; AIDS; leukemia;  
 KW carcinoma; lymphoma; antineumatic; antiarthritic; neuroprotective;  
 KW antinflammatory; antiaesthetic; antiallergic; cytostatic.

XX Unidentified.

XX WO2003055979-A2.

XX 10-JUL-2003.

XX 14-NOV-2002; 2002WO-US036496.

XX 16-NOV-2001; 2001US-0331469P.

XX 19-DEC-2001; 2001US-0340817P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;

XX WPI; 2003-505530/47.

XX Novel antibody that immunospecifically binds to a B lymphocyte stimulator  
 PT (Blys), useful for detecting and treating diseases or disorders e.g.  
 PT rheumatoid arthritis, aschma and leukemia.

XX Example 1; SEQ ID NO 1034; 394pp; English.

XX This invention relates to novel antibodies that immunospecifically bind  
 CC to B lymphocyte stimulator (Blys). The Blys gene has been mapped to  
 CC chromosome 13q34 and encodes a protein that is a member of the tumor  
 CC necrosis factor superfamily and induces both in vivo and in vitro B cell  
 CC proliferation and differentiation. Specifically, it refers to single  
 CC chain antibody molecules (scfvs) derived, preferably, from the variable  
 CC heavy CDR3 region that immunospecifically bind to a polypeptide, or  
 CC fragment thereof, of either human, murine, rat or monkey Blys. The  
 CC present invention refers to the use of such antibodies in various methods  
 CC for the detection, diagnosis and prognosis of diseases related to the

CC aberrant expression or inappropriate function of Blys or its receptor. As  
 CC such, these compositions are useful for identifying immune disorders  
 CC including myasthenia gravis and multiple sclerosis, inflammatory  
 CC disorders e.g. aschma and rheumatoid arthritis, infectious diseases such  
 CC as AIDS and proliferative disorders including leukemia, carcinoma and  
 CC lymphoma. Accordingly, they can be described as exhibiting various  
 CC activities such as antineumatic, antiarthritic, neuroprotective,  
 CC antinflammatory, antiaesthetic, antiallergic and cytostatic.  
 CC polypeptide sequence is a single chain antibody that binds Blys of the  
 CC invention. NOTE: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 259 AA;

Query Match 8.1%; Score 21; DB 7; Length 259;

Best Local Similarity 100.0%; Pred. No. 3.7e-10; Mismatches 0; Indels 0; Gaps 0;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 ESGAEVKKPGSSVKSCKASG 26  
 Db 6 ESGAEVKKPGSSVKSCKASG 26

RESULT 3  
 ABP45252  
 ID ABP45252 standard; protein; 259 AA.  
 XX  
 AC ABP45252;

DT 19-ANG-2002 (first entry)

XX Human Blys binding scfv SEQ ID 1263.

XX  
 DE  
 XX  
 KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;  
 KW tumor necrosis factor; B cell proliferation; B cell differentiation;  
 KW immunosuppressive; immunostimulant; immunomodulatory; antineumatic;  
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;  
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;  
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.

XX Homo sapiens.

XX WO200202641-A1.

XX 10-JAN-2002.

XX 15-JUN-2001; 2001WO-US019110.

XX 16-JUN-2000; 2000US-0212210P.

XX 17-OCT-2000; 2000US-0240816P.

XX 16-MAR-2001; 2001US-0276248P.

XX 21-MAR-2001; 2001US-0277379P.

XX 25-MAY-2001; 2001US-0293499P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.

XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;

XX WPI; 2002-114799/15.

XX Antibodies against B lymphocyte stimulating polypeptides, useful for the  
 PT diagnosis and treatment of cancers and immune disorders.  
 PT  
 XX Claim 1; Page 1909-1911; 3148pp; English.

XX This invention describes novel antibodies that immunospecifically bind to  
 CC B lymphocyte stimulator (Blys) polypeptides. Blys is a member of the  
 CC tumor necrosis factor (TNF) super family and induces B cell  
 CC proliferation and differentiation. The antibodies of the invention have  
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,  
 CC antineumatic and antiAIDS activity and can be used in vaccines to

CC inhibit the expression and activity of Blys. The antibodies bind to Blys  
 CC and so may be used to detect and quantitate the presence of Blys in  
 CC biological samples and may be used in this way to diagnose disease  
 CC associated with aberrant expression of Blys. They may also be  
 CC administered to treat diseases associated with aberrant Blys expression  
 CC and activity such as cancer, immune, and autoimmune disorders and  
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,  
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and  
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent  
 CC the antibodies and fragments of the antibodies described in the method of  
 CC the invention  
 CC  
 SQ Sequence 259 AA;  
 XX  
 XX  
 Query Match 7.7%; Score 20; DB 5; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 SGAEEKKPGSSVYVSCKASG 26  
 DB 7 SGAEEKKPGSSVYVSCKASG 26  
 RESULT 4  
 ABP45264  
 ID ABP45264 standard; protein; 259 AA.  
 XX  
 AC ABP45264;  
 XX  
 DT 19-AUG-2002 (first entry)  
 XX  
 DE Human Blys binding scFv SEQ ID 1275.  
 XX  
 KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;  
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;  
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;  
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;  
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;  
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200202641-A1.  
 XX  
 PD 10-JAN-2002.  
 XX  
 PF 15-JUN-2001; 2001WO-US019110.  
 XX  
 PR 16-JUN-2000; 2000US-0212210P.  
 PR 17-OCT-2000; 2000US-0240816P.  
 PR 16-MAR-2001; 2001US-0276248P.  
 PR 21-MAR-2001; 2001US-0277379P.  
 PR 25-MAY-2001; 2001US-0293499P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.  
 XX  
 PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;  
 XX  
 DR WPI; 2002-114799/15.  
 XX  
 PT Antibodies against B lymphocyte stimulating polypeptides, useful for the  
 PT diagnosis and treatment of cancers and immune disorders.  
 XX  
 PS Claim 1; Page 1924-1925; 3148pp; English.  
 XX  
 CC This invention describes novel antibodies that immunospecifically bind to  
 CC B lymphocyte stimulator (Blys) polypeptides. Blys is a member of the  
 CC tumour necrosis factor (TNF) super family and induces B cell  
 CC proliferation and differentiation. The antibodies of the invention have  
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,  
 CC antirheumatic and antiAIDS activity and can be used in vaccines to  
 CC inhibit the expression and activity of Blys. The antibodies bind to Blys

CC and so may be used to detect and quantitate the presence of Blys in  
 CC biological samples and may be used in this way to diagnose disease  
 CC associated with aberrant expression of Blys. They may also be  
 CC administered to treat diseases associated with aberrant Blys expression  
 CC and activity such as cancer, immune, and autoimmune disorders and  
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,  
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and  
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent  
 CC the antibodies and fragments of the antibodies described in the method of  
 CC the invention  
 CC  
 SQ Sequence 259 AA;  
 XX  
 XX  
 Query Match 7.7%; Score 20; DB 5; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 SGAEEKKPGSSVYVSCKASG 26  
 DB 7 SGAEEKKPGSSVYVSCKASG 26  
 RESULT 5  
 ABP45236  
 ID ABP45236 standard; protein; 259 AA.  
 XX  
 AC ABP45236;  
 XX  
 DT 19-AUG-2002 (first entry)  
 XX  
 DE Human Blys binding scFv SEQ ID 1247.  
 XX  
 KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;  
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;  
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;  
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;  
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;  
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200202641-A1.  
 XX  
 PD 10-JAN-2002.  
 XX  
 PF 15-JUN-2001; 2001WO-US019110.  
 XX  
 PR 16-JUN-2000; 2000US-0212210P.  
 PR 17-OCT-2000; 2000US-0240816P.  
 PR 16-MAR-2001; 2001US-0276248P.  
 PR 21-MAR-2001; 2001US-0277379P.  
 PR 25-MAY-2001; 2001US-0293499P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.  
 XX  
 PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;  
 XX  
 DR WPI; 2002-114799/15.  
 XX  
 PT Antibodies against B lymphocyte stimulating polypeptides, useful for the  
 PT diagnosis and treatment of cancers and immune disorders.  
 XX  
 PS Claim 1; Page 1890-1891; 3148pp; English.  
 XX  
 CC This invention describes novel antibodies that immunospecifically bind to  
 CC B lymphocyte stimulator (Blys) polypeptides. Blys is a member of the  
 CC tumour necrosis factor (TNF) super family and induces B cell  
 CC proliferation and differentiation. The antibodies of the invention have  
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,  
 CC antirheumatic and antiAIDS activity and can be used in vaccines to  
 CC inhibit the expression and activity of Blys. The antibodies bind to Blys  
 CC and so may be used to detect and quantitate the presence of Blys in

CC biological samples and may be used in this way to diagnose disease  
CC associated with aberrant expression of Blys. They may also be  
CC administered to treat diseases associated with aberrant Blys expression  
CC and activity such as cancer, immune, and autoimmune disorders and  
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,  
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and  
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP4728 represent  
CC the antibodies and fragments of the antibodies described in the method of  
CC the invention  
XX  
SQ Sequence 259 AA;  
XX  
Query Match 7.7%; Score 20; DB 5; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
Qy 7 SGAEVKKPGSSVVKVSCKASG 26  
Db 7 SGAEVKKPGSSVVKVSCKASG 26

RESULT 6  
ID ABP45032 standard; protein; 259 AA.  
XX  
AC ABP45032;  
XX  
DT 19-AUG-2002 (first entry)  
XX  
DE Human Blys binding scFv SEQ ID 1043.  
XX  
KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;  
KW tumour necrosis factor; B cell proliferation; B cell differentiation;  
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;  
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;  
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;  
KW common variable immunodeficiency; acquired immunodeficiency syndrome.  
XX  
OS Homo sapiens.  
XX  
PN WO200202641-A1.  
XX  
PD 10-JAN-2002.  
XX  
PF 15-JUN-2001; 2001WO-US019110.  
XX  
PR 16-JUN-2000; 2000US-0212210P.  
PR 17-OCT-2000; 2000US-0240816P.  
PR 16-MAR-2001; 2001US-0276248P.  
PR 21-MAR-2001; 2001US-0277379P.  
PR 25-MAY-2001; 2001US-0293499P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.  
XX  
PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;  
XX  
DR WPI; 2002-114799/15.  
XX  
PT Antibodies against B lymphocyte stimulating polypeptides, useful for the  
PT diagnosis and treatment of cancers and immune disorders.  
XX  
Claim 1; Page 1646-1648; 3148pp; English.  
XX  
CC This invention describes novel antibodies that immunospecifically bind to  
CC B lymphocyte stimulator (Blys) polypeptides. Blys is a member of the  
CC tumour necrosis factor (TNF) super family and induces B cell  
CC proliferation and differentiation. The antibodies of the invention have  
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,  
CC antirheumatic and antiAIDS activity and can be used in vaccines to  
CC inhibit the expression and activity of Blys. The antibodies bind to Blys  
CC and so may be used to detect and quantitate the presence of Blys in  
CC biological samples and may be used in this way to diagnose disease

CC associated with aberrant expression of Blys. They may also be  
CC administered to treat diseases associated with aberrant Blys expression  
CC and activity such as cancer, immune, and autoimmune disorders and  
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,  
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and  
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP4728 represent  
CC the antibodies and fragments of the antibodies described in the method of  
CC the invention  
XX  
SQ Sequence 259 AA;  
XX  
Query Match 7.7%; Score 20; DB 5; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
Qy 7 SGAEVKKPGSSVVKVSCKASG 26  
Db 7 SGAEVKKPGSSVVKVSCKASG 26

RESULT 7  
ID ABP45259 standard; protein; 259 AA.  
XX  
AC ABP45259;  
XX  
DT 19-AUG-2002 (first entry)  
XX  
DE Human Blys binding scFv SEQ ID 1270.  
XX  
KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;  
KW tumour necrosis factor; B cell proliferation; B cell differentiation;  
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;  
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;  
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;  
KW common variable immunodeficiency; acquired immunodeficiency syndrome.  
XX  
OS Homo sapiens.  
XX  
PN WO200202641-A1.  
XX  
PD 10-JAN-2002.  
XX  
PF 15-JUN-2001; 2001WO-US019110.  
XX  
PR 16-JUN-2000; 2000US-0212210P.  
PR 17-OCT-2000; 2000US-0240816P.  
PR 16-MAR-2001; 2001US-0276248P.  
PR 21-MAR-2001; 2001US-0277379P.  
PR 25-MAY-2001; 2001US-0293499P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.  
XX  
PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;  
XX  
DR WPI; 2002-114799/15.  
XX  
PT Antibodies against B lymphocyte stimulating polypeptides, useful for the  
PT diagnosis and treatment of cancers and immune disorders.  
XX  
Claim 1; Page 1918-1919; 3148pp; English.  
XX  
CC This invention describes novel antibodies that immunospecifically bind to  
CC B lymphocyte stimulator (Blys) polypeptides. Blys is a member of the  
CC tumour necrosis factor (TNF) super family and induces B cell  
CC proliferation and differentiation. The antibodies of the invention have  
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,  
CC antirheumatic and antiAIDS activity and can be used in vaccines to  
CC inhibit the expression and activity of Blys. The antibodies bind to Blys  
CC and so may be used to detect and quantitate the presence of Blys in  
CC biological samples and may be used in this way to diagnose disease  
CC associated with aberrant expression of Blys. They may also be

CC administered to treat diseases associated with aberrant BlyS expression  
CC and activity such as cancer, immune, and autoimmune disorders and  
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,  
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and  
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent  
CC the antibodies and fragments of the antibodies described in the method of  
CC the invention  
XX  
SQ Sequence 259 AA;  
XX  
Query Match 7.7%; Score 20; DB 5; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 7 SGAEVKKPGSSVVKVSCKASG 26  
Db 7 SGAEVKKPGSSVVKVSCKASG 26  
XX  
RESULT 8  
ABP45486  
ID ABP45486 standard; protein; 259 AA.  
XX  
AC ABP45486;  
XX  
DT 19-AUG-2002 (first entry)  
XX  
DE Human BlyS binding scFv SEQ ID 1497.  
XX  
KW BlyS; B lymphocyte stimulator; TNF superfamily; human; cytosolic;  
KW tumour necrosis factor; B cell proliferation; B cell differentiation;  
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;  
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;  
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;  
KW common variable immunodeficiency; acquired immunodeficiency syndrome.  
XX  
OS Homo sapiens.  
XX  
PN WO200202641-A1.  
XX  
PD 10-JAN-2002.  
XX  
PF 15-JUN-2001; 2001WO-US019110.  
XX  
PR 16-JUN-2000; 2000US-0212210P.  
PR 17-OCT-2000; 2000US-0240816P.  
PR 16-MAR-2001; 2001US-0276248P.  
PR 21-MAR-2001; 2001US-0277379P.  
PR 25-MAY-2001; 2001US-0293499P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.  
XX  
PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;  
XX  
DR WPI; 2002-114799/15.  
XX  
PT Antibodies against B lymphocyte stimulating polypeptides, useful for the  
PT diagnosis and treatment of cancers and immune disorders.  
XX  
PS Claim 1; Page 2189-2190; 3148pp; English.  
XX  
CC This invention describes novel antibodies that immunospecifically bind to  
CC B lymphocyte stimulator (BlyS) polypeptides. BlyS is a member of the  
CC tumour necrosis factor (TNF) super family and induces B cell  
CC proliferation and differentiation. The antibodies of the invention have  
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,  
CC antineoplastic and antiAIDS activity and can be used in vaccines to  
CC inhibit the expression and activity of BlyS. The antibodies bind to BlyS  
CC and so may be used to detect and quantitate the presence of BlyS in  
CC biological samples and may be used in this way to diagnose disease  
CC associated with aberrant expression of BlyS. They may also be  
CC administered to treat diseases associated with aberrant BlyS expression

CC and activity such as cancer, immune, and autoimmune disorders and  
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,  
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and  
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent  
CC the antibodies and fragments of the antibodies described in the method of  
CC the invention  
XX  
SQ Sequence 259 AA;  
XX  
Query Match 7.7%; Score 20; DB 5; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 7 SGAEVKKPGSSVVKVSCKASG 26  
Db 7 SGAEVKKPGSSVVKVSCKASG 26  
XX  
RESULT 9  
ADG96086  
ID ADG96086 standard; protein; 259 AA.  
XX  
AC ADG96086;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Single chain antibody that immunospecifically binds BlyS Seqid 1270.  
XX  
KW antibody; B lymphocyte stimulator; BlyS; tumour necrosis factor;  
KW B cell proliferation; differentiation; scFv; myasthenia gravis;  
KW multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;  
KW carcinoma; lymphoma; antirheumatic; antiarthritic; neuroprotective;  
KW antiinflammatory; antiasthmatic; antiallergic; cytostatic.  
XX  
OS Unidentified.  
XX  
PN WO2003055979-A2.  
XX  
PD 10-JUL-2003.  
XX  
PF 14-NOV-2002; 2002WO-US036496.  
XX  
PR 16-NOV-2001; 2001US-0331469P.  
PR 19-DEC-2001; 2001US-0340817P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;  
XX  
DR WPI; 2003-505530/47.  
XX  
PT Novel antibody that immunospecifically binds to a B lymphocyte stimulator  
PT (BlyS), useful for detecting and treating diseases or disorders e.g.  
XX  
XX rheumatoid arthritis, asthma and leukemia.  
XX  
PS Example 1; SEQ ID NO 1270; 394pp; English.  
XX  
CC This invention relates to novel antibodies that immunospecifically bind  
CC to B lymphocyte stimulator (BlyS). The BlyS gene has been mapped to  
CC chromosome 13q34 and encodes a protein that is a member of the tumour  
CC necrosis factor superfamily and induces both in vivo and in vitro B cell  
CC proliferation and differentiation. Specifically, it refers to single  
CC chain antibody molecules (scFvs) derived, preferably, from the variable  
CC heavy CDR3 region that immunospecifically bind to a polypeptide, or  
CC fragment thereof, of either human, murine, rat or monkey BlyS. The  
CC present invention refers to the use of such antibodies in various methods  
CC for the detection, diagnosis and prognosis of diseases related to the  
CC aberrant expression or inappropriate function of BlyS or its receptor. As  
CC such, these compositions are useful for identifying immune disorders  
CC including myasthenia gravis and multiple sclerosis, inflammatory  
CC disorders e.g. asthma and rheumatoid arthritis, infectious diseases such  
CC as AIDS and proliferative disorders including leukaemia, carcinoma and  
CC lymphoma. Accordingly, they can be described as exhibiting various

CC activities such as antirheumatic, antiarthritic, neuroprotective,  
CC antiinflammatory, antiasthmatic, antiallergic and cytostatic. This  
CC polypeptide sequence is a single chain antibody that binds Blys of the  
CC invention. NOTE: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 259 AA;

Query Match 7.7%; Score 20; DB 7; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYKVCCKASG 26  
Db 7 SGAEVKKPGSSVYKVCCKASG 26

RESULT 10  
ADG96091  
ID ADG96091 standard; protein; 259 AA.

XX AC ADG96091;

DT 11-MAR-2004 (first entry)

XX Single chain antibody that immunospecifically binds Blys SeqID 1275.

XX antibody: B lymphocyte stimulator; Blys; tumour necrosis factor;  
KM B cell proliferation; differentiation; scFv; myasthenia gravis;  
KM multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;  
KM carcinoma; lymphoma; antirheumatic; antiallergic; neuroprotective;  
KM antiinflammatory; antiasthmatic; antiallergic; cytostatic.

XX Unidentified.

OS NO2003055979-A2.

PN 10-JUL-2003.

PD 14-NOV-2002; 2002WO-US036496.

XX 16-NOV-2001; 2001US-0331469P.

PR 19-DEC-2001; 2001US-0340817P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;

XX WPI; 2003-505530/47.

XX Novel antibody that immunospecifically binds to a B lymphocyte stimulator  
PT (Blys), useful for detecting and treating diseases or disorders e.g.  
PT rheumatoid arthritis, asthma and leukemia.

XX Example 1; SEQ ID NO 1275; 394pp; English.

XX This invention relates to novel antibodies that immunospecifically bind  
CC to B lymphocyte stimulator (Blys). The Blys gene has been mapped to  
CC chromosome 13q34 and encodes a protein that is a member of the tumour  
CC necrosis factor superfamily and induces both in vivo and in vitro B cell  
CC proliferation and differentiation. Specifically, it refers to single  
CC chain antibody molecules (scFvs) derived, preferably, from the variable  
CC heavy CDR3 region that immunospecifically bind to a polypeptide, or  
CC fragment thereof, of either human, murine, rat or monkey Blys. The  
CC present invention refers to the use of such antibodies in various methods  
CC for the detection, diagnosis and prognosis of diseases related to the  
CC aberrant expression or inappropriate function of Blys or its receptor. As  
CC such, these compositions are useful for identifying immune disorders  
CC including myasthenia gravis and multiple sclerosis, inflammatory  
CC disorders e.g. asthma and rheumatoid arthritis, infectious diseases such  
CC as AIDS and proliferative disorders including leukaemia, carcinoma and  
CC lymphoma. Accordingly, they can be described as exhibiting various

CC activities such as antirheumatic, antiarthritic, neuroprotective,  
CC antiinflammatory, antiasthmatic, antiallergic and cytostatic. This  
CC polypeptide sequence is a single chain antibody that binds Blys of the  
CC invention. NOTE: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 259 AA;

Query Match 7.7%; Score 20; DB 7; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYKVCCKASG 26  
Db 7 SGAEVKKPGSSVYKVCCKASG 26

RESULT 11  
ADG96313  
ID ADG96313 standard; protein; 259 AA.

XX AC ADG96313;

DT 11-MAR-2004 (first entry)

XX Single chain antibody that immunospecifically binds Blys SeqID 1497.

XX antibody: B lymphocyte stimulator; Blys; tumour necrosis factor;  
KM B cell proliferation; differentiation; scFv; myasthenia gravis;  
KM multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;  
KM carcinoma; lymphoma; antirheumatic; antiallergic; neuroprotective;  
KM antiinflammatory; antiasthmatic; antiallergic; cytostatic.

XX Unidentified.

OS NO2003055979-A2.

PN 10-JUL-2003.

PD 14-NOV-2002; 2002WO-US036496.

XX 16-NOV-2001; 2001US-0331469P.

PR 19-DEC-2001; 2001US-0340817P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;

XX WPI; 2003-505530/47.

XX Novel antibody that immunospecifically binds to a B lymphocyte stimulator  
PT (Blys), useful for detecting and treating diseases or disorders e.g.  
PT rheumatoid arthritis, asthma and leukemia.

XX Example 1; SEQ ID NO 1497; 394pp; English.

XX This invention relates to novel antibodies that immunospecifically bind  
CC to B lymphocyte stimulator (Blys). The Blys gene has been mapped to  
CC chromosome 13q34 and encodes a protein that is a member of the tumour  
CC necrosis factor superfamily and induces both in vivo and in vitro B cell  
CC proliferation and differentiation. Specifically, it refers to single  
CC chain antibody molecules (scFvs) derived, preferably, from the variable  
CC heavy CDR3 region that immunospecifically bind to a polypeptide, or  
CC fragment thereof, of either human, murine, rat or monkey Blys. The  
CC present invention refers to the use of such antibodies in various methods  
CC for the detection, diagnosis and prognosis of diseases related to the  
CC aberrant expression or inappropriate function of Blys or its receptor. As  
CC such, these compositions are useful for identifying immune disorders  
CC including myasthenia gravis and multiple sclerosis, inflammatory  
CC disorders e.g. asthma and rheumatoid arthritis, infectious diseases such  
CC as AIDS and proliferative disorders including leukaemia, carcinoma and  
CC lymphoma. Accordingly, they can be described as exhibiting various



CC activities such as antirheumatic, antiarthritic, neuroprotective,  
CC antiinflammatory, antiasthmatic, antiallergic and cytostatic. This  
CC polypeptide sequence is a single chain antibody that binds Blys of the  
CC invention. NOTE: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 259 AA;

Query Match 7.7%; Score 20; DB 7; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYKVCCKASG 26  
Db 7 SGAEVKKPGSSVYKVCCKASG 26

RESULT 12

ADG96063  
ID ADG96063 standard; protein; 259 AA.

XX ADG96063;

DT 11-MAR-2004 (first entry)

XX Single chain antibody that immunospecifically binds Blys SegID 1247.

XX antibody; B lymphocyte stimulator; Blys; tumour necrosis factor;  
XX B cell proliferation; differentiation; scfv; myasthenia gravis;  
XX multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;  
XX carcinoma; lymphoma; antirheumatic; antiarthritic; neuroprotective;  
XX antiinflammatory; antiasthmatic; antiallergic; cytostatic.

OS Unidentified.

PN WO2003055979-A2.

XX 10-JUL-2003.

PD 14-NOV-2002; 2002WO-US036496.

XX 16-NOV-2001; 2001US-0331469P.

PR 19-DEC-2001; 2001US-0340817P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;

DR WPI; 2003-505530/47.

XX Novel antibody that immunospecifically binds to a B lymphocyte stimulator

PT (Blys), useful for detecting and treating diseases or disorders e.g.

PT rheumatoid arthritis, asthma and leukemia.

XX Example 1; SEQ ID NO 1247; 394pp; English.

XX This invention relates to novel antibodies that immunospecifically bind  
XX to B lymphocyte stimulator (Blys). The Blys gene has been mapped to  
XX chromosome 13q34 and encodes a protein that is a member of the tumour  
XX necrosis factor superfamily and induces both in vivo and in vitro B cell  
XX proliferation and differentiation. Specifically, it refers to single  
XX chain antibody molecules (scfvs) derived, preferably, from the variable  
XX heavy CDR3 region that immunospecifically bind to a polypeptide, or  
XX fragment thereof, of either human, murine, rat or monkey Blys. The  
XX present invention refers to the use of such antibodies in various methods  
XX for the detection, diagnosis and prognosis of diseases related to the  
XX aberrant expression or inappropriate function of Blys or its receptor. As  
XX such, these compositions are useful for identifying immune disorders  
XX including myasthenia gravis and multiple sclerosis, inflammatory  
XX disorders e.g. asthma and rheumatoid arthritis, infectious diseases such  
XX as AIDS and proliferative disorders including leukaemia, carcinoma and  
XX lymphoma. Accordingly, they can be described as exhibiting various

CC activities such as antirheumatic, antiarthritic, neuroprotective,  
CC antiinflammatory, antiasthmatic, antiallergic and cytostatic. This  
CC polypeptide sequence is a single chain antibody that binds Blys of the  
CC invention. NOTE: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 259 AA;

Query Match 7.7%; Score 20; DB 7; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYKVCCKASG 26  
Db 7 SGAEVKKPGSSVYKVCCKASG 26

RESULT 13

ADG95859  
ID ADG95859 standard; protein; 259 AA.

XX ADG95859;

DT 11-MAR-2004 (first entry)

XX Single chain antibody that immunospecifically binds Blys SegID 1043.

XX antibody; B lymphocyte stimulator; Blys; tumour necrosis factor;  
XX B cell proliferation; differentiation; scfv; myasthenia gravis;  
XX multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;  
XX carcinoma; lymphoma; antirheumatic; antiarthritic; neuroprotective;  
XX antiinflammatory; antiasthmatic; antiallergic; cytostatic.

OS Unidentified.

PN WO2003055979-A2.

XX 10-JUL-2003.

PD 14-NOV-2002; 2002WO-US036496.

XX 16-NOV-2001; 2001US-0331469P.

PR 19-DEC-2001; 2001US-0340817P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;

DR WPI; 2003-505530/47.

XX Novel antibody that immunospecifically binds to a B lymphocyte stimulator

PT (Blys), useful for detecting and treating diseases or disorders e.g.

PT rheumatoid arthritis, asthma and leukemia.

XX Example 1; SEQ ID NO 1043; 394pp; English.

XX This invention relates to novel antibodies that immunospecifically bind  
XX to B lymphocyte stimulator (Blys). The Blys gene has been mapped to  
XX chromosome 13q34 and encodes a protein that is a member of the tumour  
XX necrosis factor superfamily and induces both in vivo and in vitro B cell  
XX proliferation and differentiation. Specifically, it refers to single  
XX chain antibody molecules (scfvs) derived, preferably, from the variable  
XX heavy CDR3 region that immunospecifically bind to a polypeptide, or  
XX fragment thereof, of either human, murine, rat or monkey Blys. The  
XX present invention refers to the use of such antibodies in various methods  
XX for the detection, diagnosis and prognosis of diseases related to the  
XX aberrant expression or inappropriate function of Blys or its receptor. As  
XX such, these compositions are useful for identifying immune disorders  
XX including myasthenia gravis and multiple sclerosis, inflammatory  
XX disorders e.g. asthma and rheumatoid arthritis, infectious diseases such  
XX as AIDS and proliferative disorders including leukaemia, carcinoma and  
XX lymphoma. Accordingly, they can be described as exhibiting various

CC activities such as antirheumatic, antiarthritic, neuroprotective,  
 CC antiinflammatory, antiaesthetic, antiallergic and cytostatic. This  
 CC polypeptide sequence is a single chain antibody that binds Blys of the  
 CC invention. NOTE: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format  
 CC directly from WIPO at fcp.wipo.int/pub/published pct\_sequences.

XX Sequence 259 AA;

Query Match 7.7%; Score 20; DB 7; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYKVCCKASG 26  
 Db 7 SGAEVKKPGSSVYKVCCKASG 26

RESULT 14

ADG96079 standard; protein; 259 AA.

ADG96079;

11-MAR-2004 (first entry)

Single chain antibody that immunospecifically binds Blys SegID 1263.

antibody; B lymphocyte stimulator; Blys; tumor necrosis factor;  
 B cell proliferation; differentiation; scfv; myasthenia gravis;  
 multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;  
 carcinoma; lymphoma; antirheumatic; antiarthritic; neuroprotective;  
 antiinflammatory; antiaesthetic; antiallergic; cytostatic.

Unidentified.

WO2003055979-A2.

10-JUL-2003.

14-NOV-2002; 2002MO-US036496.

16-NOV-2001; 2001US-0331469P.

19-DEC-2001; 2001US-0340817P.

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;

WPI; 2003-505330/47.

Novel antibody that immunospecifically binds to a B lymphocyte stimulator  
 (Blys), useful for detecting and treating diseases or disorders e.g.  
 rheumatoid arthritis, asthma and leukemia.

Example 1; SEQ ID NO 1263; 394pp; English.

This invention relates to novel antibodies that immunospecifically bind  
 to B lymphocyte stimulator (Blys). The Blys gene has been mapped to  
 chromosome 13q34 and encodes a protein that is a member of the tumor  
 necrosis factor superfamily and induces both in vivo and in vitro B cell  
 proliferation and differentiation. Specifically, it refers to single  
 chain antibody molecules (scfvs) derived, preferably, from the variable  
 heavy CDR3 region that immunospecifically bind to a polypeptide, or  
 fragment thereof, of either human, murine, rat or monkey Blys. The  
 present invention refers to the use of such antibodies in various methods  
 for the detection, diagnosis and prognosis of diseases related to the  
 aberrant expression or inappropriate function of Blys or its receptor. As  
 such, these compositions are useful for identifying immune disorders  
 including myasthenia gravis and multiple sclerosis, inflammatory  
 disorders e.g. asthma and rheumatoid arthritis, infectious diseases such  
 as AIDS and proliferative disorders including leukaemia, carcinoma and  
 lymphoma. Accordingly, they can be described as exhibiting various

CC activities such as antirheumatic, antiarthritic, neuroprotective,  
 CC antiinflammatory, antiaesthetic, antiallergic and cytostatic. This  
 CC polypeptide sequence is a single chain antibody that binds Blys of the  
 CC invention. NOTE: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format  
 CC directly from WIPO at fcp.wipo.int/pub/published pct\_sequences.

XX Sequence 259 AA;

Query Match 7.7%; Score 20; DB 7; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYKVCCKASG 26  
 Db 7 SGAEVKKPGSSVYKVCCKASG 26

RESULT 15

AAU72862 standard; protein; 259 AA.

AAU72862;

26-FEB-2002 (first entry)

P4-3 single chain Fv.

Human; NKGD2; NKGD2 receptor complex; cancer; infectious disease; tumour;  
 autoimmune disease; head; neck; stomach; oesophagus; colon; liver; ovary;  
 intrahepatic bile duct; pancreas; lung; larynx; breast; uterus; cervix;  
 prostate; kidney; testis; thyroid; bladder; brain; melanoma; myeloma; Fv;  
 sarcoma; leukaemia; lymphoma; virus; bacterium; fungus; protozoan; DAPI0;  
 helminth; cytostatic; antimicrobial; immunomodulatory; 11B2D10; 6H7E7;  
 KM 8G7C10; 6E5A7; 11B2D10x4-7; 8G7C10x4-7; 6E5A7x4-7; P4-2; P4-3; P4-14;  
 KM P4-15; P5-2; P5-3; P5-9; P5-10; P5-11; P5-23; 3B10xP4-3; 3B10xP4-14;  
 KM P53 tetramerisation domain; 3B10xP5-2; 3B10xP5-23.

Homo sapiens.

WO200171005-A2.

27-SEP-2001.

26-MAR-2001; 2001WO-EP003414.

24-MAR-2000; 2000EP-00106467.

(KUEF) KUEFER P.

Kufer P, Riethmueler G, Lutterbuese R, Borschert K, Kischel R;

Mayer M, Hofmeister R;

WPI; 2002-055119/07.

N-PSDB; AAS97136.

Multi-functional polypeptides comprising binding sites that specifically  
 recognise extracellular groups of the NKGD2 receptor complex and domains  
 PT which function as receptors or ligands, useful for treating cancers and  
 PT infectious diseases.

Example 7; Fig 16; 114pp; English.

The invention relates to a multifunctional polypeptide comprising a  
 domain with a binding site that specifically recognises an extracellular  
 CC group of the NKGD2 receptor complex and a second domain which functions  
 CC as a receptor or ligand. The polypeptide and its associated  
 CC polynucleotide are used for the preparation of a pharmaceutical  
 CC composition for the treatment of cancer, infections and/or autoimmune  
 CC conditions. The cancer may be a tumour of the head and neck, stomach,  
 CC oesophagus, colon, liver, intrahepatic bile ducts, pancreas, lung,  
 CC larynx, breast, ovary, uterus, cervix, prostate, kidney, testis, thyroid,  
 CC bladder or brain, or a melanoma, myeloma, sarcoma, leukaemia or lymphoma.

CC The infectious diseases can be caused by viruses, bacteria, fungi,  
CC protozoa or helminths. The autoimmune diseases include multiple  
CC sclerosis, Grave's disease, ankylosing spondylitis, acute anterior  
CC uveitis, Goodpasture's syndrome, myasthenia gravis, insulin-dependent  
CC diabetes mellitus, rheumatoid arthritis, pemphigus vulgaris and  
CC autoimmune hepatitis. Sequences AAU72820-AAU72875 represent the NKGD  
CC receptor and the polypeptides of the invention  
XX

SQ Sequence 259 AA;

Query Match 6.9%; Score 18; DB 5; Length 259;

Best Local Similarity 100.0%; Pred. No. 1.7e-07;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 SGGGSGGGSGGGSGSEL 145

Db 119 SGGGSGGGSGGGSGSEL 136

Search completed: September 26, 2005, 13:11:54  
Job time : 162 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2005, 13:06:17 ; Search time 41 Seconds  
(without alignments)  
471.564 Million cell updates/sec

Title: US-09-936-702-3\_COPY\_244\_502  
Perfect score: 259  
Sequence: 1 QVQLLSGAEVKKRSGSVKV.....RYFGQCTRLRIKLVRGSG 259

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 384

Minimum DB seq length: 259  
Maximum DB seq length: 259

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA: \*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep: \*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep: \*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep: \*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep: \*  
5: /cgn2\_6/ptodata/1/iaa/PTUG\_COMB.pep: \*  
6: /cgn2\_6/ptodata/1/iaa/backfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	6.2	259	4 US-09-419-788-29	Sequence 29, Appl
2	10	3.9	259	4 US-09-419-788-115	Sequence 115, App
3	7	2.7	259	3 US-08-469-318-141	Sequence 141, App
4	7	2.7	259	3 US-08-468-609A-141	Sequence 141, App
5	7	2.7	259	3 US-09-216-295-5	Sequence 5, Appl
6	7	2.7	259	3 US-08-446-872A-141	Sequence 141, App
7	7	2.7	259	4 US-08-762-227A-141	Sequence 141, App
8	7	2.7	259	4 US-09-632-570-5	Sequence 5, Appl
9	7	2.7	259	4 US-09-632-575-35	Sequence 35, Appl
10	7	2.7	259	5 PCT-US95-01185-141	Sequence 141, App
11	6	2.3	259	4 US-09-079-723-236	Sequence 236, App
12	6	2.3	259	4 US-09-079-723-243	Sequence 243, App
13	6	2.3	259	4 US-09-079-723-248	Sequence 248, App
14	6	2.3	259	4 US-09-419-788-114	Sequence 114, App
15	6	2.3	259	4 US-07-795-859B-4	Sequence 4, Appl
16	5	1.9	259	1 US-08-457-616-4	Sequence 29, Appl
17	5	1.9	259	2 US-08-408-095-29	Sequence 133, App
18	5	1.9	259	3 US-08-469-318-133	Sequence 155, App
19	5	1.9	259	3 US-08-469-318-155	Sequence 155, App
20	5	1.9	259	3 US-08-468-609A-133	Sequence 133, App
21	5	1.9	259	3 US-08-468-609A-155	Sequence 155, App
22	5	1.9	259	3 US-09-082-090-2	Sequence 2, Appl
23	5	1.9	259	3 US-08-446-872A-133	Sequence 133, App
24	5	1.9	259	3 US-08-446-872A-155	Sequence 155, App
25	5	1.9	259	3 US-09-134-001C-3619	Sequence 3619, App
26	5	1.9	259	3 US-09-253-316-27	Sequence 27, Appl

28	5	1.9	259	4 US-08-762-227A-133	Sequence 133, App
29	5	1.9	259	4 US-08-762-227A-155	Sequence 155, App
30	5	1.9	259	4 US-09-252-991A-25689	Sequence 25689, A
31	5	1.9	259	4 US-09-252-991A-28573	Sequence 28573, A
32	5	1.9	259	4 US-09-252-991A-30105	Sequence 30105, A
33	5	1.9	259	4 US-09-742-693-33	Sequence 33, Appl
34	5	1.9	259	4 US-09-489-039A-9829	Sequence 9829, App
35	5	1.9	259	4 US-09-445-774-27	Sequence 229, App
36	5	1.9	259	4 US-09-976-594-229	Sequence 9, Appl
37	5	1.9	259	4 US-09-071-252-9	Sequence 476, App
38	5	1.9	259	4 US-09-602-787A-476	Sequence 33541, A
39	5	1.9	259	4 US-09-270-767-33541	Sequence 109, App
40	5	1.9	259	4 US-09-919-039-109	Sequence 15939, A
41	5	1.9	259	4 US-09-248-796A-15939	Sequence 18507, A
42	5	1.9	259	4 US-09-248-796A-18507	Sequence 20465, A
43	5	1.9	259	4 US-09-248-796A-20465	Sequence 9267, App
44	5	1.9	259	4 US-09-949-016-9267	Sequence 10093, A
45	5	1.9	259	4 US-09-949-016-10093	

ALIGNMENTS

RESULT 1  
US-09-419-788-29  
Sequence 29, Application US/09419788  
Patent No. 6825325  
GENERAL INFORMATION:  
APPLICANT: FISCHER, Rainer  
APPLICANT: SCHILLBERG, Stefan  
APPLICANT: NAHRING, Jorg  
APPLICANT: SACK, Markus  
APPLICANT: MONECKE, Michael  
APPLICANT: LIAO, Yu-Cai  
APPLICANT: SPIEGEL, Holger  
APPLICANT: ZIMMERMAN, Sabine  
APPLICANT: EMANS, Neil  
TITLE OF INVENTION: Molecular Pathogenicity Mediated Plant Disease  
FILE REFERENCE: 0147-0189P  
CURRENT APPLICATION NUMBER: US/09/419,788  
CURRENT FILING DATE: 1999-10-18  
EARLIER APPLICATION NUMBER: 98 11 9630.6 EP  
EARLIER FILING DATE: 1998-10-16  
EARLIER APPLICATION NUMBER: 66/BOM/1998 INDIA  
EARLIER FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 163  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 29  
LENGTH: 259  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURES:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic, no  
US-09-419-788-29  
Query Match 6.2%; Score 16; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 116 SGGGSGGGSGGGGS 143  
116 SGGGSGGGSGGGGS 131  
RESULT 2  
US-09-419-788-115  
Sequence 115, Application US/09419788  
Patent No. 6825325  
GENERAL INFORMATION:  
APPLICANT: FISCHER, Rainer  
APPLICANT: SCHILLBERG, Stefan

APPLICANT: NAHRING, JORG  
APPLICANT: SACK, Markus  
APPLICANT: MONECKE, Michael  
APPLICANT: LIAO, Yu-Cai  
APPLICANT: SPIEGEL, Holger  
APPLICANT: ZIMMERMAN, Sabine  
APPLICANT: EMANS, Neil  
TITLE OF INVENTION: Molecular Pathogenicity Mediated Plant Disease  
FILE REFERENCE: 0147-0189P  
CURRENT APPLICATION NUMBER: US/09/419,798  
EARLIER FILING DATE: 1998-10-16  
EARLIER APPLICATION NUMBER: 66/BOM/1998 INDIA  
EARLIER FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 163  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 115  
LENGTH: 259  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic, no  
OTHER INFORMATION: natural origin  
US-09-419-788-115

Query Match 3.9%; Score 10; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 LAMVQKPGQ 185  
DB 179 LAMVQKPGQ 188

RESULT 3  
US-08-469-318-141  
Sequence 141, Application US/08469318  
Patent No. 6022535  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion  
TITLE OF INVENTION: Protein  
NUMBER OF SEQUENCES: 196  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,318  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/446,872  
FILING DATE:  
INFORMATION FOR SEQ ID NO: 141:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 259 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-469-318-141

Query Match 2.7%; Score 7; DB 3; Length 259;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GGGSGG 136  
DB 123 GGGSGG 129

RESULT 4  
US-08-468-609A-141  
Sequence 141, Application US/08468609A  
Patent No. 6030812  
GENERAL INFORMATION:  
APPLICANT: Abrams, Mark A.  
APPLICANT: Bauer, S. C.  
APPLICANT: Braford-Goldberg, Sarah R.  
APPLICANT: Caparon, Maite H.  
APPLICANT: Easton, Alan M.  
APPLICANT: Klein, Barbara K.  
APPLICANT: McKeown, John P.  
APPLICANT: Olin, Peter O.  
APPLICANT: Paik, Kumman  
APPLICANT: Thomas, John W.  
TITLE OF INVENTION: Fusion Proteins Comprising Multiply Mutated Interleukin-3 (IL-3)  
NUMBER OF SEQUENCES: 197  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Dennis A. Bennett, G.D. Searle & Co.,  
ADDRESSER: Corporate Patent Dept.  
STREET: P. O. Box 5110  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60680

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,609A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/192,325  
FILING DATE: 14-FEB-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bennett, Dennis A.  
REGISTRATION NUMBER: 34,547  
REFERENCE/DOCKET NUMBER: C-2790/3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)737-6986  
TELEFAX: (314)737-6972  
INFORMATION FOR SEQ ID NO: 141:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 259 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-468-609A-141

Query Match 2.7%; Score 7; DB 3; Length 259;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GGGSGG 136  
DB 123 GGGSGG 129

RESULT 5  
US-09-216-295-5  
Sequence 5, Application US/09216295  
Patent No. 6268328  
GENERAL INFORMATION:  
APPLICANT: Mitchinson, Colin  
APPLICANT: Wendt, Dan J.  
TITLE OF INVENTION: No. 6268328e1 Variant EGIII-like Cellulase Compositions  
FILE REFERENCE: GC555

;; CURRENT APPLICATION NUMBER: US/09/216,295  
;; CURRENT FILING DATE: 1998-12-18  
;; NUMBER OF SEQ ID NOS: 41  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO: 5  
;; LENGTH: 259  
;; TYPE: PRT  
;; ORGANISM: Aspergillus aculeatus  
US-09-216-295-5

Query Match 2.7%; Score 7; DB 3; Length 259;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 PATSVS 157  
Db 224 PATSVS 230

## RESULT 6

US-08-446-872A-141  
; Sequence 141, Application US/08446872A  
; Patent No. 6361977  
; GENERAL INFORMATION:  
; APPLICANT: Abrams, Mark A.  
; APPLICANT: Bauer, S. C.  
; APPLICANT: Braford-Goldberg, Sarah R.  
; APPLICANT: Caparon, Maite H.  
; APPLICANT: Easton, Alan M.  
; APPLICANT: Klein, Barbara K.  
; APPLICANT: McKearn, John P.  
; APPLICANT: Olin, Peter O.  
; APPLICANT: Paik, Kumman  
; APPLICANT: Thomas, John W.  
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis  
; TITLE OF INVENTION: Fusion Protein  
; NUMBER OF SEQUENCES: 197  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,  
; ADDRESSEE: Corporate Patent Dept.  
; STREET: P. O. Box 5110  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60680  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,872A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/192,325  
; FILING DATE: 14-FEB-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bennett, Dennis A.  
; REGISTRATION NUMBER: 34,547  
; REFERENCE/DOCKET NUMBER: C-2790/1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314)737-6986  
; TELEFAX: (314)737-6972  
; INFORMATION FOR SEQ ID NO: 141:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 259 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-446-872A-141

Query Match 2.7%; Score 7; DB 3; Length 259;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GGGSGGG 136  
Db 123 GGGSGGG 129

## RESULT 7

US-08-762-227A-141  
; Sequence 141, Application US/08762227A  
; Patent No. 6436387  
; GENERAL INFORMATION:  
; APPLICANT: Abrams, Mark A.  
; APPLICANT: Bauer, S. C.  
; APPLICANT: Braford-Goldberg, Sarah R.  
; APPLICANT: Caparon, Maite H.  
; APPLICANT: Easton, Alan M.  
; APPLICANT: Klein, Barbara K.  
; APPLICANT: McKearn, John P.  
; APPLICANT: Olin, Peter O.  
; APPLICANT: Paik, Kumman  
; APPLICANT: Thomas, John W.  
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis  
; TITLE OF INVENTION: Fusion Protein  
; NUMBER OF SEQUENCES: 197  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,  
; ADDRESSEE: Corporate Patent Dept.  
; STREET: P. O. Box 5110  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60680  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/762,227A  
; FILING DATE: 09-Dec-1996  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/192,325  
; FILING DATE: 14-FEB-1994  
; APPLICATION NUMBER: US 08/446,872  
; FILING DATE: 06-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bennett, Dennis A.  
; REGISTRATION NUMBER: 34,547  
; REFERENCE/DOCKET NUMBER: C-2790/5  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (708)470-6501  
; TELEFAX: (708)470-6881  
; INFORMATION FOR SEQ ID NO: 141:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 259 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 141:  
US-08-762-227A-141

Query Match 2.7%; Score 7; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GGGSGGG 136  
Db 123 GGGSGGG 129

RESULT 8  
US-09-632-570-5  
; Sequence 5, Application US/09632570  
; Patent No. 6623949  
; GENERAL INFORMATION:  
; APPLICANT: Gualfetti, Peter  
; APPLICANT: Mitchinson, Colin  
; APPLICANT: Phillips, Jay Ian  
; TITLE OF INVENTION: No. 6623949el Variant EGIII-Like Cellulase  
; TITLE OF INVENTION: Compositions  
; FILE REFERENCE: GC631  
; CURRENT APPLICATION NUMBER: US/09/632,570  
; CURRENT FILING DATE: 2000-08-04  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Aspergillus aculeatus  
US-09-632-570-5

Query Match 2.7%; Score 7; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 PATLSVS 157  
Db 224 PATLSVS 230

RESULT 9  
US-09-632-575-35  
; Sequence 35, Application US/09632575  
; Patent No. 6635465  
; GENERAL INFORMATION:  
; APPLICANT: Gualfetti, Peter  
; APPLICANT: Mitchinson, Colin  
; APPLICANT: ROPP, Traci M.  
; TITLE OF INVENTION: Mutant EGIII Cellulase, DNA Encoding  
; TITLE OF INVENTION: Such EGIII Compositions and Methods for Obtaining Same  
; FILE REFERENCE: GC629  
; CURRENT APPLICATION NUMBER: US/09/632,575  
; CURRENT FILING DATE: 2000-08-04  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Aspergillus aculeatus  
US-09-632-575-35

Query Match 2.7%; Score 7; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 PATLSVS 157  
Db 224 PATLSVS 230

RESULT 10  
PCT-US95-01185-141  
; Sequence 141, Application PC/TUS9501185  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion  
; NUMBER OF SEQUENCES: 196  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/01185  
; FILING DATE: 02-FEB-1995  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/192325  
; FILING DATE: 14-FEB-1994  
; INFORMATION FOR SEQ ID NO: 141:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 259 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-01185-141

Query Match 2.7%; Score 7; DB 5; Length 259;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GGGSGGG 136  
Db 123 GGGSGGG 129

RESULT 11  
US-09-079-723-228  
; Sequence 228, Application US/09079723  
; Patent No. 6703362  
; GENERAL INFORMATION:  
; APPLICANT: Alvarez, Vernon L.  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; APPLICANT: Singleton, Judith  
; APPLICANT: Paterson, Catherine A.  
; APPLICANT: Cagney, Gerard M.  
; APPLICANT: Bellinka, Benjamin A.  
; APPLICANT: Carter, John M.  
; TITLE OF INVENTION: RANDOM PEPTIDES THAT BIND TO GASTRO-  
; TITLE OF INVENTION: INTESTINAL TRACT (GIT) TRANSPORT RECEPTORS AND RELATED METHODS  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/079,723  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Miarock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-219  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 228:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 259 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown



MOLECULE TYPE: peptide  
US-09-079-723-228

Query Match  
Best Local Similarity 100.0%; Score 6; DB 4; Length 259;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 LVPRGS 258  
|||||  
Db 221 LVPRGS 226

RESULT 12  
US-09-079-723-236  
Sequence 236, Application US/09079723  
Patent No. 6703362

GENERAL INFORMATION:  
APPLICANT: Alvarez, Vernon L.  
APPLICANT: O'Mahony, Daniel J.  
APPLICANT: Lambkin, Imelda J.  
APPLICANT: Singleton, Judith  
APPLICANT: Patterson, Catherine A.  
APPLICANT: Cagney, Gerard M.  
APPLICANT: Belinka, Benjamin A.  
TITLE OF INVENTION: RANDOM PEPTIDES THAT BIND TO GASTRO-  
INTESTINAL TRACT (GIT) TRANSPORT RECEPTORS AND RELATED METHODS  
NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/079,723  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Mierock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-219  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 236:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 259 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-09-079-723-236

Query Match  
Best Local Similarity 100.0%; Score 6; DB 4; Length 259;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 LVPRGS 258  
|||||  
Db 221 LVPRGS 226

RESULT 13  
US-09-079-723-243  
Sequence 243, Application US/09079723

Patent No. 6703362

GENERAL INFORMATION:  
APPLICANT: Alvarez, Vernon L.  
APPLICANT: O'Mahony, Daniel J.  
APPLICANT: Lambkin, Imelda J.  
APPLICANT: Singleton, Judith  
APPLICANT: Patterson, Catherine A.  
APPLICANT: Cagney, Gerard M.  
APPLICANT: Belinka, Benjamin A.  
TITLE OF INVENTION: RANDOM PEPTIDES THAT BIND TO GASTRO-  
INTESTINAL TRACT (GIT) TRANSPORT RECEPTORS AND RELATED METHODS  
NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/079,723  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Mierock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-219  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 243:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 259 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-09-079-723-243

Query Match  
Best Local Similarity 100.0%; Score 6; DB 4; Length 259;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 LVPRGS 258  
|||||  
Db 221 LVPRGS 226

RESULT 14  
US-09-079-723-248  
Sequence 248, Application US/09079723  
Patent No. 6703362

GENERAL INFORMATION:  
APPLICANT: Alvarez, Vernon L.  
APPLICANT: O'Mahony, Daniel J.  
APPLICANT: Lambkin, Imelda J.  
APPLICANT: Singleton, Judith  
APPLICANT: Patterson, Catherine A.  
APPLICANT: Cagney, Gerard M.  
APPLICANT: Belinka, Benjamin A.  
TITLE OF INVENTION: RANDOM PEPTIDES THAT BIND TO GASTRO-  
INTESTINAL TRACT (GIT) TRANSPORT RECEPTORS AND RELATED METHODS  
NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/079,723  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Mastrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-219  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 248:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 259 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-09-079-723-248

Query Match 2.3%; Score 6; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 LVPRGS 258  
Db 221 LVPRGS 226

RESULT 15  
US-09-419-788-114  
Sequence 114, Application US/09419788  
Patent No. 6825325  
GENERAL INFORMATION:  
APPLICANT: FISCHER, Rainer  
APPLICANT: SCHILLBERG, Stefan  
APPLICANT: NAHRING, Jorg  
APPLICANT: SACK, Markus  
APPLICANT: MONECKE, Michael  
APPLICANT: LIAO, Yu-Cai  
APPLICANT: SPIEGEL, Holger  
APPLICANT: ZIMMERMAN, Sabine  
APPLICANT: EMANS, Neil  
TITLE OF INVENTION: Molecular Pathogenicicid Mediated Plant Disease  
TITLE OF INVENTION: Resistance  
FILE REFERENCE: 0147-0189P  
CURRENT APPLICATION NUMBER: US/09/419,788  
CURRENT FILING DATE: 1999-10-18  
EARLIER APPLICATION NUMBER: 98 11 9630.6 EP  
EARLIER FILING DATE: 1998-10-16  
EARLIER APPLICATION NUMBER: 66/BOM/1998 INDIA  
EARLIER FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 163  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 114  
LENGTH: 259  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic, no  
US-09-419-788-114

Query Match 2.3%; Score 6; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 118 WGGCTL 123  
Db 113 WGGCTL 118

Search completed: September 26, 2005, 13:16:25  
Job time : 42 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 26, 2005, 13:14:58 ; Search time 168 Seconds  
(without alignments)  
627.497 Million cell updates/sec

Title: US-09-936-702-3\_COPY\_244\_502

Perfect score: 259  
Sequence: 1 QVQLRSGAEVKKPKGSSVKY.....RTYFGQGTURLIKLVRGSG 259

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1826554 seqs, 407025358 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3391

Minimum DB seq length: 259

Maximum DB seq length: 259

Post-processing: Listing first 45 summaries

## Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubppaa/PCTR\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubppaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubppaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubppaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/2/pubppaa/US10E\_PUBCOMB.pep:\*  
18: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/2/pubppaa/US11\_PUBCOMB.pep:\*  
20: /cgn2\_6/ptodata/2/pubppaa/US11\_NEW\_PUB.pep:\*  
21: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep:\*  
22: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	9.7	259	US-10-863-729-18	Sequence 18, Appl
2	21	8.1	259	US-09-880-748-1034	Sequence 1034, Ap
3	21	8.1	259	US-10-293-418-1034	Sequence 1034, Ap
4	20	7.7	259	US-09-880-748-1043	Sequence 1043, Ap
5	20	7.7	259	US-09-880-748-1247	Sequence 1247, Ap
6	20	7.7	259	US-09-880-748-1263	Sequence 1263, Ap
7	20	7.7	259	US-09-880-748-1270	Sequence 1270, Ap
8	20	7.7	259	US-09-880-748-1275	Sequence 1275, Ap
9	20	7.7	259	US-09-880-748-1457	Sequence 1457, Ap
10	20	7.7	259	US-10-293-418-1043	Sequence 1043, Ap
11	20	7.7	259	US-10-293-418-1247	Sequence 1247, Ap

12	20	7.7	259	15	US-10-293-418-1263	Sequence 1263, Ap
13	20	7.7	259	15	US-10-293-418-1270	Sequence 1270, Ap
14	20	7.7	259	15	US-10-293-418-1275	Sequence 1275, Ap
15	20	7.7	259	15	US-10-293-418-1497	Sequence 1497, Ap
16	18	6.9	259	15	US-10-293-656-53	Sequence 53, Appl
17	16	6.2	259	10	US-09-880-748-864	Sequence 864, Appl
18	16	6.2	259	10	US-09-880-748-973	Sequence 973, Appl
19	16	6.2	259	10	US-09-880-748-1025	Sequence 1025, Ap
20	16	6.2	259	10	US-09-880-748-1032	Sequence 1032, Ap
21	16	6.2	259	10	US-09-880-748-1041	Sequence 1041, Ap
22	16	6.2	259	10	US-09-880-748-1262	Sequence 1262, Ap
23	16	6.2	259	10	US-09-880-748-1277	Sequence 1277, Ap
24	16	6.2	259	10	US-09-880-748-1356	Sequence 1356, Ap
25	16	6.2	259	10	US-09-880-748-1452	Sequence 1452, Ap
26	16	6.2	259	10	US-09-880-748-1481	Sequence 1481, Ap
27	16	6.2	259	10	US-09-880-748-1485	Sequence 1485, Ap
28	16	6.2	259	10	US-09-880-748-1552	Sequence 1552, Ap
29	16	6.2	259	10	US-09-880-748-1567	Sequence 1567, Ap
30	16	6.2	259	10	US-09-880-748-1664	Sequence 1664, Ap
31	16	6.2	259	15	US-10-293-418-864	Sequence 864, Appl
32	16	6.2	259	15	US-10-293-418-973	Sequence 973, Appl
33	16	6.2	259	15	US-10-293-418-1025	Sequence 1025, Ap
34	16	6.2	259	15	US-10-293-418-1032	Sequence 1032, Ap
35	16	6.2	259	15	US-10-293-418-1041	Sequence 1041, Ap
36	16	6.2	259	15	US-10-293-418-1262	Sequence 1262, Ap
37	16	6.2	259	15	US-10-293-418-1277	Sequence 1277, Ap
38	16	6.2	259	15	US-10-293-418-1356	Sequence 1356, Ap
39	16	6.2	259	15	US-10-293-418-1452	Sequence 1452, Ap
40	16	6.2	259	15	US-10-293-418-1481	Sequence 1481, Ap
41	16	6.2	259	15	US-10-293-418-1485	Sequence 1485, Ap
42	16	6.2	259	15	US-10-293-418-1552	Sequence 1552, Ap
43	16	6.2	259	15	US-10-293-418-1567	Sequence 1567, Ap
44	16	6.2	259	15	US-10-293-418-1664	Sequence 1664, Ap
45	15	5.8	259	14	US-10-207-655-13	Sequence 13, Appl

## ALIGNMENTS

RESULT 1  
US-10-863-729-18  
; Sequence 18, Application US/10863729  
; Publication No. US20050013819A1  
; GENERAL INFORMATION:  
; APPLICANT: Kinch, Michael S.  
; TITLE OF INVENTION: USE OF EPH4 AND MODULATOR OF EPH4 FOR  
; FILE REFERENCE: 10271-117-999  
; CURRENT APPLICATION NUMBER: US/10/863, 729  
; CURRENT FILING DATE: 2004-06-07  
; PRIOR APPLICATION NUMBER: 60/476, 909  
; PRIOR FILING DATE: 2003-06-06  
; PRIOR APPLICATION NUMBER: 60/503, 356  
; PRIOR FILING DATE: 2003-09-16  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 18  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; OTHER INFORMATION: Mature scFv (EA44) sequence with FLAG & HIS Tags  
US-10-863-729-18

Query Match 9.7%; Score 25; DB 17; Length 259;  
Best Local Similarity 100.0%; Pred. No. 3.7e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 LAWYQKPGQAPRLIYGASTRATG 200  
DB 164 LAWYQKPGQAPRLIYGASTRATG 188

```
RESULT 2
US-09-880-748-1034
; Sequence 1034, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1034
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1034

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 259;
Pred. No. 1e-09; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;

QY      6  ESGAEVKKPGSSVVKSCASG 26
Db      6  ESGAEVKKPGSSVVKSCASG 26

RESULT 3
US-10-293-418-1034
; Sequence 1034, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; PRIOR FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1034
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1034

Query Match
Best Local Similarity 100.0%; Score 21; DB 15; Length 259;
Pred. No. 1e-09; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;

QY      6  ESGAEVKKPGSSVVKSCASG 26
Db      6  ESGAEVKKPGSSVVKSCASG 26
```

```
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6  ESGAEVKKPGSSVVKSCASG 26
Db      6  ESGAEVKKPGSSVVKSCASG 26

RESULT 4
US-09-880-748-1043
; Sequence 1043, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1043
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1043

Query Match
Best Local Similarity 100.0%; Score 20; DB 10; Length 259;
Pred. No. 7.5e-09; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;

QY      7  SGAEVKKPGSSVVKSCASG 26
Db      7  SGAEVKKPGSSVVKSCASG 26

RESULT 5
US-09-880-748-1247
; Sequence 1247, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1247
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1247

Query Match
Best Local Similarity 100.0%; Score 20; DB 10; Length 259;
Pred. No. 7.7%; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;

QY      7  SGAEVKKPGSSVVKSCASG 26
Db      7  SGAEVKKPGSSVVKSCASG 26
```

Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYVSCKASG 26  
Db 7 SGAEVKKPGSSVYVSCKASG 26

## RESULT 6

US-09-880-748-1263  
; Sequence 1263, Application US/09880748  
; Publication No. US20030059937A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OR INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PFS523  
; CURRENT APPLICATION NUMBER: US/09/880,748  
; CURRENT FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/212,210  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/276,248  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 3239  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1263  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-748-1263

Query Match 7.7%; Score 20; DB 10; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYVSCKASG 26  
Db 7 SGAEVKKPGSSVYVSCKASG 26

## RESULT 7

US-09-880-748-1270  
; Sequence 1270, Application US/09880748  
; Publication No. US20030059937A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OR INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PFS523  
; CURRENT APPLICATION NUMBER: US/09/880,748  
; CURRENT FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/212,210  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/276,248  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 3239  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1270  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-748-1270

Query Match 7.7%; Score 20; DB 10; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYVSCKASG 26  
Db 7 SGAEVKKPGSSVYVSCKASG 26

## RESULT 8

US-09-880-748-1275  
; Sequence 1275, Application US/09880748  
; Publication No. US20030059937A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OR INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PFS523  
; CURRENT APPLICATION NUMBER: US/09/880,748  
; CURRENT FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/212,210  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/276,248  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 3239  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1275  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-748-1275

Query Match 7.7%; Score 20; DB 10; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYVSCKASG 26  
Db 7 SGAEVKKPGSSVYVSCKASG 26

## RESULT 9

US-09-880-748-1497  
; Sequence 1497, Application US/09880748  
; Publication No. US20030059937A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OR INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PFS523  
; CURRENT APPLICATION NUMBER: US/09/880,748  
; CURRENT FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/212,210  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/276,248  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 3239  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1497  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-748-1497

Query Match 7.7%; Score 20; DB 10; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVSVKSCASG 26  
|||||  
DB 7 SGAEVKKPGSSVSVKSCASG 26

RESULT 10

US-10-293-418-1043  
; Sequence 1043, Application US/10293418  
; Publication No. US20030223996A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PF523P2  
; CURRENT APPLICATION NUMBER: US/10/293,418  
; PRIOR FILING DATE: 2002-11-27  
; PRIOR APPLICATION NUMBER: 60/331,469  
; PRIOR FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: 60/340,817  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 09/880,748  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/276,248  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/212,210  
; PRIOR FILING DATE: 2000-06-16  
; NUMBER OF SEQ ID NOS: 3247  
; SEQ ID NO 1043  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-293-418-1043

Query Match 7.7%; Score 20; DB 15; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVSVKSCASG 26  
|||||  
DB 7 SGAEVKKPGSSVSVKSCASG 26

RESULT 11

US-10-293-418-1247  
; Sequence 1247, Application US/10293418  
; Publication No. US20030223996A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PF523P2  
; CURRENT APPLICATION NUMBER: US/10/293,418  
; PRIOR FILING DATE: 2002-11-27  
; PRIOR APPLICATION NUMBER: 60/331,469  
; PRIOR FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: 60/340,817  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 09/880,748  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/276,248

; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/212,210  
; PRIOR FILING DATE: 2000-06-16  
; NUMBER OF SEQ ID NOS: 3247  
; SEQ ID NO 1247  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-293-418-1247

Query Match 7.7%; Score 20; DB 15; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVSVKSCASG 26  
|||||  
DB 7 SGAEVKKPGSSVSVKSCASG 26

RESULT 12

US-10-293-418-1263  
; Sequence 1263, Application US/10293418  
; Publication No. US20030223996A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PF523P2  
; CURRENT APPLICATION NUMBER: US/10/293,418  
; PRIOR FILING DATE: 2002-11-27  
; PRIOR APPLICATION NUMBER: 60/331,469  
; PRIOR FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: 60/340,817  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 09/880,748  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/276,248  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/212,210  
; PRIOR FILING DATE: 2000-06-16  
; NUMBER OF SEQ ID NOS: 3247  
; SEQ ID NO 1263  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-293-418-1263

Query Match 7.7%; Score 20; DB 15; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVSVKSCASG 26  
|||||  
DB 7 SGAEVKKPGSSVSVKSCASG 26

RESULT 13

US-10-293-418-1270  
; Sequence 1270, Application US/10293418  
; Publication No. US20030223996A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PF523P2  
; CURRENT APPLICATION NUMBER: US/10/293,418  
; PRIOR FILING DATE: 2002-11-27

```
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1270
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1270
```

```
Query Match      7.7%; Score 20; DB 15; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      7 SGAEVKKPGSSVVKSCKASG 26
DB      7 SGAEVKKPGSSVVKSCKASG 26
```

```
RESULT 14
US-10-293-418-1275
; Sequence 1275, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: P5523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; PRIOR FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1275
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1275
```

```
Query Match      7.7%; Score 20; DB 15; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      7 SGAEVKKPGSSVVKSCKASG 26
DB      7 SGAEVKKPGSSVVKSCKASG 26
```

```
RESULT 15
US-10-293-418-1497
; Sequence 1497, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: P5523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; PRIOR FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1497
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1497
```

```
Query Match      7.7%; Score 20; DB 15; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      7 SGAEVKKPGSSVVKSCKASG 26
DB      7 SGAEVKKPGSSVVKSCKASG 26
```

Search completed: September 26, 2005, 13:28:36  
Job time : 168 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: September 26, 2005, 13:05:27 ; Search time 40 Seconds

(without alignments)  
623.004 Million cell updates/sec

Title: US-09-936-702-3\_COPY\_244\_502

Perfect score: 259  
Sequence: 1 QVQLPSGAEVKKRSGSSVKY.....RYTFGGCTRLRLIKLVRGSG 259

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 603

Minimum DB seq length: 259

Maximum DB seq length: 259

Post-processing: listing first 45 summaries

Database :

PIR\_79:.\*  
1: p1r1:.\*  
2: p1r2:.\*  
3: p1r3:.\*  
4: p1r4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	2.3	259	2 T15126	hypothetical prote
2	6	2.3	259	2 C84165	hypothetical prote
3	6	2.3	259	2 F86475	hypothetical prote
4	5	1.9	259	1 C47069	acetolactate decar
5	5	1.9	259	1 IOHUL	insulin-like growt
6	5	1.9	259	1 S46286	RNA-binding protei
7	5	1.9	259	1 D49343	cis-1,2-dihydro-1,
8	5	1.9	259	2 F83479	probable short-cha
9	5	1.9	259	2 B55217	cis-1,2-dihydro-1,
10	5	1.9	259	2 B83838	oxidoreductase BH1
11	5	1.9	259	2 C75410	methionine aminope
12	5	1.9	259	2 T78147	class II histocomp
13	5	1.9	259	2 T48119	P-glycoprotein - C
14	5	1.9	259	2 AC0358	probable pili chap
15	5	1.9	259	2 D75372	probable nitrogen
16	5	1.9	259	2 AF1847	hypothetical prote
17	5	1.9	259	2 G97397	ubiquinone/menaqui
18	5	1.9	259	2 T73185	probable gntR-fami
19	5	1.9	259	2 AC3100	ABC transporter, m
20	5	1.9	259	2 G98186	hypothetical prote
21	5	1.9	259	2 C95939	probable spermidin
22	5	1.9	259	2 B64514	hypothetical prote
23	5	1.9	259	2 AF0999	conserved hypotet
24	5	1.9	259	2 A55274	3-oxo-5alpha-stero
25	5	1.9	259	2 S65744	steroid 5alpha-red
26	5	1.9	259	2 F81877	hypothetical prote
27	5	1.9	259	2 B84886	probable beta-expa
28	5	1.9	259	2 T50657	beta-expansin limp
29	5	1.9	259	2 T13260	hypothetical prote

30	5	1.9	259	2 D86685	prophage p11 prote
31	5	1.9	259	2 C86797	prophage p13 prote
32	5	1.9	259	2 C86757	prophage p12 prote
33	5	1.9	259	2 T23782	hypothetical prote
34	5	1.9	259	2 G96673	hypothetical prote
35	5	1.9	259	2 E87537	hypothetical prote
36	5	1.9	259	2 T51745	RNA helicase RH16
37	5	1.9	259	2 D84343	TRK potassium upla
38	5	1.9	259	2 T17889	glycine-rich prote
39	5	1.9	259	2 B69811	hypothetical prote
40	5	1.9	259	2 JE0326	peroxin Pex1lp iso
41	5	1.9	259	2 C97664	hypothetical prote
42	4	1.5	259	1 DEBCSP	sorbitol-6-phospha
43	4	1.5	259	1 YXBYT	thymidylate syntha
44	4	1.5	259	1 XYBPT2	site-specific DNA-
45	4	1.5	259	1 XYBPT4	site-specific DNA-

#### ALIGNMENTS

##### RESULT 1

T15126  
hypothetical protein T20B6.3 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C:Accession: T15126

R:Beck, C.; Wamsley, P.

submitted to the EMBL Data Library, April 1997

A:Description: The sequence of C. elegans cosmid T20B6.

A:Reference number: Z18297

A:Accession: T15126

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-259 <BRC>

A:Cross-references: UNIPROT:O02049; EMBL:AF000193; NID:g1946986; PID:g1946988; PIDN:AA85;

A:Experimental source: strain Bristol N2; clone T20B6

C:Genetics:

A:Gene: CESP:T20B6.3

A:Map position: 3

A:Inserts: 9/2; 231/1

C:Superfamily: Phaeolus glycine-rich protein 1.0

Query Match 2.3%; Score 6; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 GSGGGG 136  
|||

Db 80 GSGGGG 85

##### RESULT 2

C84165  
hypothetical protein Vng0049h [imported] - *Halobacterium* sp.

C:Species: *Halobacterium* sp. NRC-1

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C:Accession: C84165

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berguist, B.; Pan, M.; Shukla, H.D.; Lasky, S.

; Leitbauer, B.; Keller, K.; Cruz, R.; Dawson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor

Jung, K.H.; Alam, M.; Freltas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li

A:Title: Genome sequence of *Halobacterium* species NRC-1.

A:Reference number: A84160; WUID:20504483; PMID:11016950

A:Accession: C84165

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-259 <STO>

A:Cross-references: UNIPROT:Q9HSM3; GB:AE004437; NID:g105796699; PIDN:AA818687.1; GSPDB:GT

C:Genetics:

A:Gene: VNG0049H

Query Match 2.3%; Score 6; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GSSSVK 20  
|||  
Db 134 GSSVKV 139

RESULT 3  
F86475  
hypothetical protein F12A4.2 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C/Accession: F86475  
R/Theologas, A.; Becker, J.R.; Palm, C.J.; Federapfel, N.A.; Kaul, S.; White, O.; Alonso,  
Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Hultzer, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lurov, J.S.; Maitl, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: F86475  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-259 <STO>  
A/Cross-references: UNIPROT:Q9C6P8; GB:A005172; NID:g10092190; PIDN:AA612608.1; GSTDB:C  
C/Genetics:  
A/Map position: 1

Query Match 2.3%; Score 6; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 SSGSGG 211  
|||||  
Db 107 SSGSGG 112

RESULT 4  
C47069  
acetylactate decarboxylase (EC 4.1.1.5) - Klebsiella terrigena  
C/Species: Klebsiella terrigena  
C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C/Accession: C47069  
R/Blomqvist, K.; Nikkola, M.; Lehtovaara, P.; Suikko, M.L.; Airaksinen, U.; Straby, K.B.  
J. Bacteriol. 175, 1392-1404, 1993  
A/Title: Characterization of the genes of the 2,3-butanediol operons from Klebsiella ter  
A/Reference number: A47069; MUID:93186707; PMID:8444801  
A/Accession: C47069  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-259 <BIO>  
A/Cross-references: UNIPROT:Q04518; GB:I04507; NID:g149170; PIDN:AAA25054.1; PID:g149171  
A/Experimental source: VTT-E-74023  
A/Note: sequence extracted from NCBI backbone (NCBI:126766, NCBI:126767)  
C/Superfamily: acetylactate decarboxylase  
C/Keyword: carbon-carbon lyase; carboxy-lyase

Query Match 1.9%; Score 5; DB 1; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 GYVEG 102  
|||||  
Db 41 GYVEG 45

RESULT 5

IOH1  
insulin-like growth factor-binding protein 1 precursor [validated] - human  
N/Alternate names: IGF-binding protein; IGF-BP25; insulin-like growth factor 25k binding  
C/Species: Homo sapiens (man)  
C/Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 09-Jul-2004  
C/Accession: A31867; S01159; A40155; A40925; A42176; S43207; S03559; S01333; A27694; S014  
R/Brinkman, A.; Groffen, C.A.H.; Kortleve, D.J.; Drop, S.L.S.  
Biochem. Biophys. Res. Commun. 157, 898-907, 1988  
A/Title: Organization of the gene encoding the insulin-like growth factor binding protei  
A/Reference number: A31867; MUID:89087480; PMID:2849945  
A/Accession: A31867  
A/Molecule type: DNA  
A/Residues: 1-259 <BR>  
A/Cross-references: UNIPROT:P08833; GB:M23592; GB:M23593; GB:M23594; GB:M23595; NID:g1844  
R/Brinkman, A.; Groffen, C.; Kortleve, D.J.; Geurts van Kessel, A.; Drop, S.L.S.  
EMBO J. 7, 2417-2423, 1988  
A/Title: Isolation and characterization of a cDNA encoding the low molecular weight [insu]  
A/Reference number: S01159; MUID:89052654; PMID:2461294  
A/Accession: S01159  
A/Molecule type: mRNA  
A/Residues: 1-252, 'W', 254-259 <BR2>  
A/Cross-references: EMBL:Y00856; NID:g32608; PIDN:CAA68770.1; PID:g32609  
A/Note: 169-Thr was also found  
R/Cubbage, M.L.; Suwanichkul, A.; Powell, D.R.  
Mol. Endocrinol. 3, 846-851, 1989  
A/Title: Structure of the human chromosome1 gene for the 25 kilodalton insulin-like growt  
A/Reference number: A40155; MUID:89330502; PMID:2474129  
A/Accession: A40155  
A/Molecule type: DNA  
A/Residues: 1-259 <CUB>  
A/Cross-references: GB:M59316; NID:g184809; PIDN:AAA52793.1; PID:g184810  
R/Lee, Y.L.; Hinz, R.B.; James, P.M.; Lee, P.D.K.; Shively, J.E.; Powell, D.R.  
Mol. Endocrinol. 2, 404-411, 1988  
A/Title: Insulin-like growth factor (IGF) binding protein complementary deoxyribonucleic  
m those of the IGF-I and IGF-II receptors.  
A/Reference number: A40925; MUID:88334540; PMID:2458522  
A/Accession: A40925  
A/Molecule type: mRNA  
A/Residues: 1-259 <LEE>  
A/Cross-references: GB:M31145; NID:g183117; PIDN:AAA52542.1; PID:g183118  
A/Note: part of this sequence, including the amino end of the mature protein, was confir  
R/Ehrenborg, E.; Larsson, C.; Stern, I.; Janson, M.; Powell, D.R.; Luthman, H.  
Genomics 12, 497-502, 1992  
A/Title: Contiguous localization of the genes encoding human insulin-like growth factor I  
A/Reference number: A42176; MUID:92217971; PMID:1373120  
A/Accession: A42176  
A/Molecule type: DNA  
A/Residues: 174-252, 'W', 254-259 <EHR>  
A/Note: sequence extracted from NCBI backbone (NCBI:95701, NCBI:95702)  
R/Luthman, H.; Seoderling-Barros, J.; Persson, B.; Engberg, C.; Stern, I.; Lake, M.; Frar  
Povoa, G.; Hall, K.; Joernvall, H.  
Eur. J. Biochem. 180, 259-265, 1989  
A/Title: Human insulin-like growth-factor-binding protein. Low-molecular-mass form: prote  
A/Reference number: S03559; MUID:89170723; PMID:2466655  
A/Accession: S43207  
A/Molecule type: mRNA  
A/Residues: 141-259 <LUT1>  
A/Cross-references: EMBL:X15002; NID:g33013; PIDN:CAA33110.1; PID:g33014  
A/Accession: S03559  
A/Molecule type: protein  
A/Residues: 26-259 <LUT2>  
A/Note: 228-Met was also found  
R/Julkinen, M.; Koistinen, R.; Aalto-Setälä, K.; Seppälä, M.; Jaenne, O.A.; Kontula,  
FEBS Lett. 236, 295-302, 1988  
A/Title: Primary structure of human insulin-like growth factor-binding protein/placental  
A/Reference number: S01333; MUID:88312985; PMID:2457513  
A/Accession: S01333  
A/Molecule type: mRNA  
A/Residues: 1-252, 'W', 254-259 <JUL>  
A/Cross-references: EMBL:X13405; NID:g35574; PIDN:CAA31771.1; PID:g35575  
R/Brewer, M.T.; Stetler, G.L.; Squires, C.H.; Squires, R.C.; Busby, W.H.; Clemmons, D.R.  
Biochem. Biophys. Res. Commun. 152, 1289-1297, 1988  
A/Title: Cloning, characterization, and expression of a human insulin-like growth factor

A;Reference number: A27694; MUID:88240345; PMID:2454104  
 A;Accession: A27694  
 A;Molecule type: mRNA  
 A;Residues: 1-54; 'PAAVAVRCAPCLWAPR', 72-212, 'Q', 214-259 <BRE>  
 A;Cross-references: GB:W20841; NID:G183113; PIDN:AA52540.1; PID:G183114  
 A;Note: the authors translated the codon CAG for residue 213 as His; the nucleotide sequ  
 R;Grundmann, U.; Nerlich, C.; Bohm, H.; Rehn, T.  
 Nucleic Acids Res. 16, 8711, 1988  
 A;Title: Cloning of cDNA encoding human placental protein 12 (PP12): binding protein for  
 A;Reference number: S01415; MUID:88335621; PMID:3419931  
 A;Accession: S01415  
 A;Molecule type: mRNA  
 A;Residues: 1-259 <GNU>  
 A;Cross-references: EMBL:X12385; NID:G35571; PIDN:CAA30942.1; PID:G35572  
 R;Busby Jr., W.H.; Klapper, D.G.; Clemmons, D.R.  
 J. Biol. Chem. 263, 14203-14210, 1988  
 A;Title: Purification of a 31,000-Dalton insulin-like growth factor binding protein from  
 A;Reference number: A30804; MUID:89008261; PMID:2971653  
 A;Accession: A30804  
 A;Molecule type: protein  
 A;Residues: 26-53 <BUS>  
 R;Bell, S.C.; Keyte, J.W.  
 Endocrinology 123, 1202-1204, 1988  
 A;Title: N-terminal amino acid sequence of human pregnancy-associated endometrial alpha-  
 weight IGF binding proteins.  
 A;Reference number: A61099; MUID:88283527; PMID:2456201  
 A;Accession: A61099  
 A;Molecule type: protein  
 A;Residues: 26-40, 'P', 42 <BEL>  
 C;Comment: This protein binds insulin-like growth factor I (somatomedin C) and insulin-I  
 C;Genetics: This protein is not glycosylated.  
 A;Gene: GDB:IGFBP1; IBP1  
 A;Cross-references: GDB:120075; OMIM:146730  
 A;Map position: 7p13-7p12  
 A;Intons: 117/1; 173/3; 216/3  
 C;Superfamily: Insulin-like growth factor binding protein; thyroglobulin type I repeat H  
 C;Keywords: phosphoprotein  
 F;1-25/Domain: signal sequence #status predicted <SIG>  
 F;16-259/Product: insulin-like growth factor-binding protein 1 #status experimental <MUT  
 F;176-251/Domain: thyroglobulin type I repeat homology <THY1>  
 F;246-248/Region: cell attachment (R-G-D) motif

Query Match 1.9%; Score 5; DB 1; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 LSCRA 168  
 Db 82 LSCRA 86

RESULT 6  
 S46286  
 RNA-binding protein - wood tobacco  
 C;Species: Nicotiana sylvestris (wood tobacco)  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C;Accession: S46286  
 R;Hirose, T.; Sugita, M.; Sugitara, M.  
 Mol. Gen. Genet. 244, 360-366, 1994  
 A;Title: Characterization of a cDNA encoding a novel type of RNA-binding protein in toba  
 A;Reference number: S46286; MUID:94359458; PMID:8078461  
 A;Accession: S46286  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-259 <HIR>  
 A;Cross-references: UNIPROT:Q40436; GB:D26182; NID:G575607; PIDN:BA05170.1; PID:G624925  
 C;Superfamily: wood tobacco RNA-binding protein; ribonucleoprotein repeat homology  
 F;41-108/Domain: ribonucleoprotein repeat homology <RML>

Query Match 1.9%; Score 5; DB 1; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 GSGCG 136  
 Db 129 GSGCG 133

RESULT 7  
 D49343  
 cis-1,2-dihydro-1,2-dihydroxynaphthalene dehydrogenase (EC 1.3.1.29) - Pseudomonas sp.  
 C;Species: Pseudomonas sp.  
 C;Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 26-Aug-1999  
 C;Accession: D49343; S27635  
 R;Denome, S.A.; Stanley, D.C.; Olson, E.S.; Young, K.D.  
 J. Bacteriol. 175, 6890-6901, 1993  
 A;Title: Metabolism of dibenzochlorophene and naphthalene in Pseudomonas strains: complete  
 A;Reference number: A49343; MUID:94042852; PMID:8226631  
 A;Accession: D49343  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-259 <DEN>  
 A;Cross-references: GB:M60405  
 C;Genetics:

A;Gene: dox  
 C;Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
 C;Keywords: oxidoreductase  
 F;5-183/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 1.9%; Score 5; DB 2; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 GSGGS 133  
 Db 147 GSGGS 151

RESULT 8  
 F83479  
 probable short-chain dehydrogenase PAL330 [imported] - Pseudomonas aeruginosa (strain PA  
 C;Species: Pseudomonas aeruginosa  
 C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
 C;Accession: F83479  
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lilm,  
 .; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
 A;Reference number: A82950; MUID:20437337; PMID:10984043  
 A;Accession: F83479  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-259 <STO>  
 A;Cross-references: UNIPROT:Q91414; GB:AE004562; GB:AE004091; NID:G9947263; PIDN:AA60471;  
 A;Experimental source: strain PA01  
 C;Genetics:

A;Gene: PAL330  
 C;Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 1.9%; Score 5; DB 2; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 EADBG 107  
 Db 229 EADBG 233

RESULT 9  
 E55117  
 cis-1,2-dihydro-1,2-dihydroxynaphthalene dehydrogenase (EC 1.3.1.29) - Pseudomonas putid  
 N;Alternate names: polycyclic aromatic hydrocarbon dihydrodiol dehydrogenase  
 C;Species: Pseudomonas putida  
 C;Date: 05-May-1995 #sequence\_revision 05-May-1995 #text\_change 09-Jul-2004

C/Accession: E55217  
R/Takizawa, N.; Kaida, N.; Torigoe, S.; Moritani, T.; Sawada, T.; Satoh, S.; Kiyohara, H.  
J. Bacteriol. 176, 2444-2449, 1994  
A/Title: Identification and characterization of genes encoding polycyclic aromatic hydrocarbons  
A/Reference number: A55217; MUID:94209249; PMID:8157615  
A/Accession: E55217  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-259 <TRK>  
A/Cross-references: UNIPROT:Q08669; GB:AB004059; GB:D16629; NID:G2189972; PIDN:BAA20393.  
C/Genetics:  
A/Genes: pahB  
C/Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
F/5-183/Domain: short-chain alcohol dehydrogenase homology <SADR>  
  
Query Match 1.9%; Score 5; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 129 GGGGS 133  
Db 147 GGGGS 151  
  
RESULT 10  
B83838  
oxidoreductase B81506 [imported] - Bacillus halodurans (strain C-125)  
C/Species: Bacillus halodurans  
C/Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
C/Accession: B83838  
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A/Reference number: A83650; MUID:20512582; PMID:11058132  
A/Accession: B83838  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-259 <STO>  
A/Cross-references: UNIPROT:Q9KCR3; GB:AP001512; GB:BA000004; NID:G10174030; PIDN:BA052  
A/Experimental source: strain C-125  
C/Genetics:  
A/Genes: B81506  
C/Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
  
Query Match 1.9%; Score 5; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 186 APRIL 190  
Db 244 APRIL 248  
  
RESULT 11  
C75410  
mechanine aminopeptidase - Deinococcus radiodurans (strain R1)  
C/Species: Deinococcus radiodurans  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C/Accession: C75410  
R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathavan, J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A/Reference number: A75250; MUID:20036896; PMID:10567266  
A/Accession: C75410  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-259 <WHI>  
A/Cross-references: UNIPROT:Q9RUS1; GB:AE001978; GB:AE000513; NID:G6459059; PIDN:AAF1088  
A/Experimental source: strain R1  
C/Genetics:

A/Genes: DR1311  
A/Map position: 1  
C/Superfamily: Escherichia coli methionyl aminopeptidase  
  
Query Match 1.9%; Score 5; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 70 ITADK 74  
Db 233 ITADK 237  
  
RESULT 12  
T28147  
class II histocompatibility antigen M beta chain 2 - chicken  
C/Species: Gallus gallus (chicken)  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000  
C/Accession: T28147  
R/Milne, S.; Kaufman, J.; Beck, S.  
submitted to the EMBL Data Library, May 1998  
A/Description: DNA sequencing and analysis of the chicken major histocompatibility comple  
A/Reference number: Z20475  
A/Accession: T28147  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-259 <MTL>  
A/Cross-references: EMBL:AL023516; PIDN:CAA18967.1  
A/Experimental source: clone CB12  
C/Genetics:  
A/Genes: BMD2  
A/Map position: 116  
A/Intons: 22/1; 116/1; 208/1; 246/1; 256/3  
C/Superfamily: class II histocompatibility antigen; immunoglobulin homology  
  
Query Match 1.9%; Score 5; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 24 ASGDT 28  
Db 160 ASGDT 164  
  
RESULT 13  
I48119  
P-glycoprotein - Chinese hamster (fragment)  
C/Species: Citreolus griseus (Chinese hamster)  
C/Date: 04-Sep-1997 #sequence\_revision 04-Sep-1997 #text\_change 02-Feb-2001  
R/Ng, W.F.; Sarangi, F.; Zastawny, R.L.; Veinot-Drebot, L.; Ling, V.  
Mol. Cell. Biol. 9, 1224-1232, 1989  
A/Title: Identification of members of the P-glycoprotein multigene family.  
A/Reference number: I48119; MUID:89261726; PMID:2566508  
A/Accession: I48119  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-259 <RES>  
A/Cross-references: GB:M25792; NID:G576810; PIDN:AA53439.1; PID:G576812  
C/Genetics:  
A/Genes: pgp3  
A/Intons: 66/3; 142/3; 191/3  
C/Superfamily: multidrug resistance protein; ATP-binding cassette homology  
C/Keywords: ATP; nucleotide binding; P-loop  
F/25-228/Domain: ATP-binding cassette homology <ABC2>  
F/42-49/Region: nucleotide-binding motif A (P-loop)  
  
Query Match 1.9%; Score 5; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 VOLUME 6  
Db 2 VOLUME 6

Db 52 VQLE 56

# RESULT 14

AC0358  
probable pilli chaperone protein YP02944 [imported] - Yersinia pestis (strain CO92)  
C/Species: Yersinia pestis  
C/Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C/Accession: AC0358  
R/Parikhll, J.; Wren, B.W.; Thomson, N.R.; Tltball, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
Nature 413, 523-527, 2001  
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A/Reference number: AB0001; MUID:21470413; PMID:11586360  
A/Accession: AC0358  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-259 <KUR>  
A/Cross-references: UNIPROT:Q82CN7; GB:AL590842; PIDN:CAC92190.1; PID:G15980902; GSPDB:  
C/Genetics:  
A/Gene: YP02944  
C/Superfamily: chaperone protein papd

Query Match 1.9%; Score 5; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 LTISS 220  
Db 18 LTISS 22

# RESULT 15

D75372  
probable nitrogen regulator - Deinococcus radiodurans (strain R1)  
C/Species: Deinococcus radiodurans  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C/Accession: D75372  
R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
S.; Shen, M.; Vamathevan, J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Me  
Science 286, 1571-1577, 1999  
A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A/Reference number: A75250; MUID:20036896; PMID:10567266  
A/Accession: D75372  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-259 <WHI>  
A/Cross-references: UNIPROT:Q9RTV7; GB:AE002007; GB:AE000513; NID:g6459402; PIDN:AAF115  
C/Genetics:  
A/Gene: DR1646  
A/Map position: 1  
C/Superfamily: regulatory protein fnr; cAMP receptor protein cyclic nucleotide-binding d

Query Match 1.9%; Score 5; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 GQGTU 123  
Db 132 GQGTU 136

Search completed: September 26, 2005, 13:15:37  
Job time : 41 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2005, 13:04:41 ; Search time 169 seconds  
(without alignments)  
784.784 Million cell updates/sec

Title: US-09-936-702-3\_COPY\_244\_502  
Perfect score: 259  
Sequence: 1 QVQLSGAEVKKPKGSSVRY.....RYTFGGTRLEIKLVERGSG 259

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1613378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2784

Minimum DB seq length: 259  
Maximum DB seq length: 259

Post-processing: Listing first 45 summaries

Database :

UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	2.3	259	2 Q8ZT71	Q8ZT71 pyrobaculum
2	6	2.3	259	2 Q9HSM3	Q9HSM3 halobacteri
3	6	2.3	259	2 Q6QHA6	Q6QHA6 homo sapien
4	6	2.3	259	2 Q02049	Q02049 caenorhabdi
5	6	2.3	259	2 Q9V6A4	Q9V6A4 drosophila
6	6	2.3	259	2 Q6QHA4	Q6QHA4 oryza sativ
7	6	2.3	259	2 Q6ZGS2	Q6ZGS2 oryza sativ
8	6	2.3	259	2 Q7XZHO	Q7XZHO oryza sativ
9	6	2.3	259	2 Q851D7	Q851D7 oryza sativ
10	6	2.3	259	2 Q9C8P8	Q9C8P8 arabidopsis
11	6	2.3	259	2 Q9F862	Q9F862 triticum ae
12	6	2.3	259	2 Q9F875	Q9F875 triticum ae
13	6	2.3	259	2 Q66FN3	Q66FN3 yerisina ps
14	6	2.3	259	2 Q6S9V1	Q6S9V1 leucospira
15	6	2.3	259	2 Q7ARN6	Q7ARN6 yerisina ps
16	6	2.3	259	2 Q7B1Z4	Q7B1Z4 yerisina ps
17	6	2.3	259	2 Q7B1Y2	Q7B1Y2 yerisina ps
18	6	2.3	259	2 Q83W92	Q83W92 escherichia
19	6	2.3	259	2 Q7ARKO	Q7ARKO yerisina ps
20	6	2.3	259	2 P74994	P74994 escherichia
21	6	2.3	259	2 Q6MOL1	Q6MOL1 bdellovibri
22	6	2.3	259	2 Q741Z8	Q741Z8 mycobacteri
23	6	2.3	259	1 ALDC_KLEETE	ALDC_KLEETE klebsiella
24	6	2.3	259	1 EXB4_ARATVH	EXB4_ARATVH arabidopsis
25	6	2.3	259	1 IBP1_HUMAN	IBP1_HUMAN homo sapien
26	6	2.3	259	1 NAHB_PSEPU	NAHB_PSEPU pseudomonas
27	6	2.3	259	1 PYRK_BACAN	PYRK_BACAN bacillus an
28	6	2.3	259	1 PYRK_BACCR	PYRK_BACCR bacillus an
29	6	2.3	259	1 S5A1_HUMAN	S5A1_HUMAN homo sapien
30	6	2.3	259	1 S5A1_RAT	S5A1_RAT rattus norv
31	6	2.3	259	1 SURE_BACTN	SURE_BACTN bacteroides

32	5	1.9	259	1 UBIE_AGRTS	Q8UH5 agrobacteri
33	5	1.9	259	1 YB14_NEIMA	Q9JUV9 neisseria m
34	5	1.9	259	1 YZ38_METJA	Q60293 methanococc
35	5	1.9	259	2 Q648P5	Q648P5 uncultured
36	5	1.9	259	2 Q8PY11	Q8PY11 methanosarc
37	5	1.9	259	2 Q8ZUQ2	Q8ZUQ2 pyrobaculum
38	5	1.9	259	2 Q9HNV8	Q9HNV8 halobacteri
39	5	1.9	259	2 Q96V53	Q96V53 pleurotus o
40	5	1.9	259	2 Q6CKL3	Q6CKL3 kluyveromyc
41	5	1.9	259	2 Q6CKG1	Q6CKG1 kluyveromyc
42	5	1.9	259	2 Q6T413	Q6T413 leptospira
43	5	1.9	259	2 Q7S5M0	Q7S5M0 neurospora
44	5	1.9	259	2 Q7S669	Q7S669 neurospora
45	5	1.9	259	2 Q96011	Q96011 homo sapien

#### ALIGNMENTS

```
RESULT 1
08ZT71      PRELIMINARY;      PRT;      259 AA.
AC  08ZT71;
DT  01-MAR-2002 (TEMBLrel. 20, Created)
DT  01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT  01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE  Hypothetical protein PAE3407.
GN  OrderedlocusNames=PAE3407;
OS  Pyrobaculum aerophilum.
OC  Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC  Thermoproteaceae; Pyrobaculum.
OX  NCBI_TaxID=13773;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=IM2 / ATCC 51768 / DSM 7523.
RX  MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA  Fitz-Gibbon S.T., Ladhner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA  Miller J.H.;
RT  "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT  aerophilum.";
RL  Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR  EMBL; AE009928; AAL64892.1; -.
KW  Complete proteome.
SQ  SEQUENCE 259 AA; 27170 MW; 726890475F69CC00 CRC64;

Query Match      2.3%; Score 6; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  54 ILDVAN 59
Db  231 ILDVAN 236

RESULT 2
09HSM3      PRELIMINARY;      PRT;      259 AA.
AC  09HSM3;
DT  01-MAR-2001 (TEMBLrel. 16, Created)
DT  01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT  01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE  Vng0049h.
GN  OrderedlocusNames=VNG0049H;
OS  Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC  Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;
OC  Halobacteriaceae; Halobacterium.
OX  NCBI_TaxID=64091;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=20504463; PubMed=11016950; DOI=10.1073/pnas.190337797;
RA  Ng W.-V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA  Shukla H.D., Lasky S.R., Baling N.S., Thorsson V., Sbrogna J.,
RA  Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
```

RA Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
 RA Maddocks D.G., Jabloncki P.E., Krebs M.P., Angevine C.M., Dale H.,  
 RA Isenbarger T.A., Peck R.F., Pohlschoder M., Spudich J.L., Jung K.-H.,  
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
 RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.,  
 RT "Genome sequence of Halobacterium species NRC-1."  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
 DR EMBL; AS004973; MAG18687.1; -.  
 DR PIR; C84165; C84165.  
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.  
 DR InterPro; IPR006342; FkM\_mtfase.  
 DR InterPro; IPR000051; SAM\_Bind.  
 DR TrEMBL; TIGR01444; FkM\_fam. 1.  
 KW Complete proteome.  
 SQ SEQUENCE 259 AA; 28842 MW; 6D147C0A0CC0784C CRC64;

Query Match 2.3%; Score 6; DB 2; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 4e+03;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GSSGVV 20  
 DB 134 GSSGVV 139

RESULT 3  
 O60HA6 PRELIMINARY; PRT; 259 AA.  
 AC O60HA6;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Myonectrin (Fragment).  
 GN Name=MYNN;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=20334280; PubMed=10873615; DOI=10.1006/dbpc.2000.2862;  
 RA Alliel P.M., Seddigi N., Goudou D., Cifuentes-Diaz C., Romero N.,  
 RA Velasco E., Rieger F., Perin J.P.,  
 RT "Myonectrin, a novel member of the BTB/POZ-zinc finger family highly  
 expressed in human muscle."  
 RL Biochem. Biophys. Res. Commun. 273:385-391(2000).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX PubMed=14694499;  
 RA Cifuentes-Diaz C., Bitoun M., Goudou D., Seddigi N., Romero N.,  
 RA Rieger F., Perin J.P., Alliel P.M.;  
 RT "Neuromuscular expression of the BTB/POZ and zinc finger protein  
 myonectrin."  
 RL Muscle Nerve 29:59-65(2004).  
 DR EMBL; AY541761; AA587377.1; -.  
 DR GO; GO:0005515; F:Protein binding; IEA.  
 DR InterPro; IPR000210; BTB\_POZ.  
 DR Pfam; PR00651; BTB. 1.  
 DR SMART; SM00225; BTB. 1.  
 DR PROSITE; PSS0097; BTB. 1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 259 AA; 28992 MW; 556F938E5133EB9F CRC64;

Query Match 2.3%; Score 6; DB 2; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 4e+03;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 SELELT 148  
 DB 202 SELELT 207

RESULT 4  
 O02049 PRELIMINARY; PRT; 259 AA.  
 AC O02049;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein T20B6.3.  
 GN Name=T20B6.3; ORFNames=T20B6.3;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=9089613; PubMed=9851916;  
 RG Wormbase Consortium;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 investigating biology. The C. elegans Sequencing Consortium."  
 RL Science 282:2012-2016(1998).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Beck C., Wamsley P.;  
 RT "The sequence of C. elegans cosmid T20B6."  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 RN (4)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RG Wormbase Consortium;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF000193; AA52890.1; -.  
 DR PIR; T15126; T15126.  
 DR Wormbase; WBGene00020599; T20B6.3.  
 DR WormPep; T20B6.3; CE13778.  
 DR InterPro; IPR002952; Eggshell.  
 DR PRINTS; PR01228; EGGSHL.  
 KW Hypothetical protein.  
 SQ SEQUENCE 259 AA; 23054 MW; C5F1690D2CE04CAF CRC64;

Query Match 2.3%; Score 6; DB 2; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 4e+03;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 GSGGGG 136  
 DB 80 GSGGGG 85

RESULT 5  
 O9V6A4 PRELIMINARY; PRT; 259 AA.  
 AC O9V6A4;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE CG30042-PA.  
 GN ORFNames=CG30042;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN (1)  
 RP SEQUENCE FROM N.A.



RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amandides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfankuch C., Baldwin D.,  
 RA Ballew R.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolashkov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burlis K.C., Busam D.A., Butler H., Cadieu L.B., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,  
 RA Palazzolo M., Plitman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skipski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodruff M., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri U.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*";  
 RL Science 287:2185-2195(2000).  
 RN  
 RP  
 RP  
 RX MEDLINE=22426065; PubMed=12537566;  
 RA Celinker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frishe E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
 RA Paclob J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
 RL *melanogaster* euchromatic genome sequence";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN  
 RN  
 RP  
 RP  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskaas R.,  
 RA Patel S., Frishe E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celinker S.E.;  
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
 RL a genomic perspective";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN  
 RN  
 RP  
 RP  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Baccourt B.R., Celinker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RL systematic review";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).

RN [5]  
 RP SEQUENCE FROM N.A.  
 RG FLYBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RG FLYBase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AEO03822; AAF58524.2; -;  
 DR FLYBase; FBgn0050042; CG30042.  
 DR FLYBase; FBgn0050045; CG30045.  
 DR GO; GO:0042302; F:structural constituent of cuticle; IEA.  
 DR InterPro; IPR005479; Cphp\_synth\_L\_D2.  
 DR InterPro; IPR000618; Insect cuticle.  
 DR Pfam; PF00379; Chitin bind\_4; 1.  
 DR PRINTS; PR00947; CUTICLE.  
 DR PROSITE; PS00867; CPASASE\_2; UNKNOWN\_1.  
 DR PROSITE; PS00233; CUTICLE; UNKNOWN\_1.  
 SQ SEQUENCE 259 AA; 26957 MW; ED2A2B6758BF1FD0 CRC64;

Query Match 2.3%; Score 6; DB 2; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 4e+03;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 129 GGGGSG 134  
 DB 105 GGGGSG 110

## RESULT 6

Q6QHA4 PRELIMINARY; PRT; 259 AA.  
 AC Q6QHA4;  
 DT 05-JUN-2004 (TEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)  
 DT 25-OCT-2004 (TEMBLrel. 28, Last annotation update)  
 DE Glutathione S-transferase.  
 GN Name=GSTU38; Synonyms=P0425F05.18;  
 OS *Oryza sativa* (Japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae;  
 OC Eriactoidae; Oryzaceae; Oryza.  
 OX NCBI\_Taxid=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA PubMed=15063639; DOI=10.1007/s00438-004-1006-8;  
 RA Soranzo N., Sati Goria M., Mizzi L., De Toma G., Firova C.;  
 RT "Organisation and structural evolution of the rice glutathione S-  
 RT transferase gene family";  
 RL Mol. Genet. Genomics 271:511-521(2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC  
 RT clone:PO425F05.";  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY541763; AA593256.1; -;  
 DR EMBL; AP003569; BAD37467.1; -;  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR InterPro; IPR010987; GST\_C\_1like.  
 DR InterPro; IPR004045; GST\_Nterm.  
 DR Pfam; PF02798; GST\_N; 1.  
 KW Transferase.  
 SQ SEQUENCE 259 AA; 27478 MW; 526A5BDP7C0F3D84 CRC64;

Query Match 2.3%; Score 6; DB 2; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 4e+03;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 185 QAPRL 190  
 DB 101 QAPRL 106

```

RESULT 7
Q6ZGS2 PRELIMINARY; PRT; 259 AA.
AC Q6ZGS2;
DT 05-JUL-2004 (TEMBLrel. 27, Created)
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DE Plant disease resistance polypeptide-like.
Name=OJ1148.D05.10;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriarthroideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasakura T., Matsumoto T., Yamamoto K.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF004118; BAD07691.1; -.
KW Polypeptide.
SQ SEQUENCE 259 AA; 27424 MW; F380F493462F69DA CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 132 GSGGCG 137
Db 60 GSGGCG 65

RESULT 8
Q7XZH0 PRELIMINARY; PRT; 259 AA.
AC Q7XZH0;
DT 01-OCT-2003 (TEMBLrel. 25, Created)
DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)
DT 25-OCT-2004 (TEMBLrel. 28, Last annotation update)
DE Hypothetical protein OSJNB0033J23.5 (Hypothetical protein
OSJNB0028F23.3).
GN Name=OSJNB0033J23.5; Synonym=OSJNB0028F23.3;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriarthroideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Teltzin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zisemann V., Heiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Uteback T.T., Feldlyum T.V.,
RA Yang Q.Q., Haas B.J., Sub B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Teltzin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zisemann V., Heiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Uteback T.T., Feldlyum T.V.,
RA Yang Q.Q., Haas B.J., Sub B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNB0028F23 genomic sequence.";
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Buell R.;

```

```

RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC137507; AAP73847.1; -.
DR EMBL; AC135595; AAT77902.1; -.
DR Gramene; Q7XZH0; -.
KW Hypothetical protein.
SQ SEQUENCE 259 AA; 28114 MW; 1805F80A7CEBD524 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 130 GSGGCG 135
Db 158 GSGGCG 163

RESULT 9
Q85ID7 PRELIMINARY; PRT; 259 AA.
AC Q85ID7;
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Hypothetical protein OSJNB0021P10.5.
GN Name=OSJNB0021P10.5;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriarthroideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Teltzin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zisemann V., Heiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Uteback T.T., Feldlyum T.V.,
RA Yang Q.Q., Haas B.J., Sub B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC123974; AA019373.1; -.
DR Gramene; Q85ID7; -.
DR InterPro; IPR010847; Hml.
DR Pfam; PF07320; Hml; 1.
KW Hypothetical protein.
SQ SEQUENCE 259 AA; 27972 MW; 03F79FD4DCE2E701 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 126 VTSGGC 131
Db 138 VTSGGC 143

RESULT 10
Q9C8P8 PRELIMINARY; PRT; 259 AA.
AC Q9C8P8;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE Helix-loop-helix protein 1A, putative; 28707-26892 (At1g35460/F12A4_2)
DE (putative bHLH transcription factor).
GN Name=F12A4.2;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.

```

```

OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Iln X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Meitz R., Ronning C.M., Koo H., Fujii C.Y., Uterback T.R.,
RA Barnstead M.B., Bowman C.J., White O., Nieman W.C., Fraser C.M.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Meyers M.C., Shin P., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Shin P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA TISSUE=Flower;
RX MEDLINE=2598051; PubMed=12679534; DOI=10.1093/molbev/meg088;
RA Heim M.A., Jakoby M., Werber M., Martin C., Weisshaar B., Bailey P.C.;
RT "The basic helix-loop-helix transcription factor family in plants: a
RT genome-wide study of protein structure and functional diversity.";
RL Mol. Biol. Evol. 20:735-747(2003).
RN [6]
RP SEQUENCE FROM N.A.
RA TISSUE=Flower;
RC Jakoby M.J., Heim M.A., Bailey P., Martin C., Weisshaar B.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
DR EMBL; AC023064; AAG52112.1; -.
DR EMBL; AY093794; AAL0410.1; -.
DR EMBL; AF411791; AAL06481.1; -.
DR EMBL; AF488612; AAM10958.1; -.
DR PIR; F86475; F86475.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SMO0353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
SQ SEQUENCE 259 AA; 28526 MW; D4B99C758208C1D4 CRC64;

Query Match 2.3%; Score 6; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 SGGSG 211
DB 107 SGGSG 112

```

```

DT 01-MAR-2004 (TREMBlrel. 26, last annotation update)
DE Gamma-gliadin (Fragment).
GN Name=GAG56D;
OS Triticum aestivum subsp. spelta.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticaceae; Triticum.
OX NCBI_TaxID=58933;
RN [1]
RP SEQUENCE FROM N.A.
RA von Bueren M.;
RT "Polymorphisms in two homeologous gamma-gliadin genes and the
RT evolution of cultivated wheat.";
RL Genet. Resour. Crop Evol. 0:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RA von Bueren M.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ389674; CAC11087.1; -.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR001376; Gliadin.
DR Pfam; PF00234; Tryp_alpha_amyl; 1.
DR PRINTS; PR00208; GLIADGLUTEN.
DR SMART; SMO0499; AAI; 1.
FT NON_TER 1 259
SQ SEQUENCE 259 AA; 29961 MW; 89A3073443C4191A CRC64;

Query Match 2.3%; Score 6; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 GQCTLV 124
DB 215 GQCTLV 220

RESULT 12
ID Q9FS75 PRELIMINARY; PRT; 259 AA.
AC Q9FS75;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 26, last sequence update)
DE Gamma-gliadin (Fragment).
GN Name=GAG56D;
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticaceae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA von Bueren M.;
RT "Polymorphisms in two homeologous gamma-gliadin genes and the
RT evolution of cultivated wheat.";
RL Genet. Resour. Crop Evol. 0:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RA von Bueren M.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ389674; CAC11057.1; -.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR001376; Gliadin.
DR Pfam; PF00234; Tryp_alpha_amyl; 1.
DR PRINTS; PR00208; GLIADGLUTEN.
DR SMART; SMO0499; AAI; 1.

```

```

FT  NON_TER      1      1
FT  NON_TER      259    259
SQ  SEQUENCE      259 AA; 30020 MW; BECF13EB3B0812B9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 259;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      119 GGGTLV 124
      |||||
DB      216 GGGTLV 221

RESULT 13
Q66FN3
ID  Q66FN3      PRELIMINARY;      PRT;      259 AA.
AC  Q66FN3;
DT  25-OCT-2004 (TReMBLrel. 28, Last Created)
DT  25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DE  Putative IS100 transposase.
GN  ORFNames=YPTB0303, YPTB1133, YPTB1586, YPTB1809, YPTB3876;
OS  Yersinia pseudotuberculosis IP 32953.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC  Enterobacteriaceae; Yersinia.
OX  NCBI_TaxID=273123;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=IP 32953;
RX  PubMed=1535858;
RA  Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA  Regala W.M., Georgescu A.M., Vergez L.W., Land M.L., Motin L.V.,
RA  Brubaker R.R., Fowler J., Hinebusch B.J., Marceau M., Medigue C.,
RA  Simonet M., Cheneal-Francois V., Souza B., Dacheux D., Elliott J.M.,
RA  Derbise A., Hauser L.J., Garcia E.;
RT  "Insights into the genome evolution of Yersinia pestis through whole
RT  genome comparison with Yersinia pseudotuberculosis.";
RL  Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR  EMBL; BX936398; CAH19543.1; -.
DR  EMBL; BX936398; CAH20373.1; -.
DR  EMBL; BX936398; CAH20825.1; -.
DR  EMBL; BX936398; CAH21048.1; -.
DR  EMBL; BX936398; CAH23114.1; -.
DR  InterPro; IPR003593; AAA_ATPase.
DR  InterPro; IPR001270; Chaprin_c1pA/B.
DR  InterPro; IPR002611; IsetB_ATPbind.
DR  Pfam; PFO1695; IsetB.1.
DR  PRINTS; PR00300; CLPROTEASEA.
DR  SMART; SM00382; AAA.1.
KW  ATP-binding.
SQ  SEQUENCE      259 AA; 29295 MW; E72FD9033B80E379 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 259;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      186 APRLLI 191
      |||||
DB      164 APRLLI 169

RESULT 14
Q6S9V1
ID  Q6S9V1      PRELIMINARY;      PRT;      259 AA.
AC  Q6S9V1;
DT  05-JUL-2004 (TReMBLrel. 27, Created)
DT  05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT  05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE  LipL32 (Fragment).
GN  Name=LipL32;
OS  Leptospira borgpetersenii.
OC  Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OC  NCBI_TaxID=174;

```

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Mus 127;
RX  PubMed=15090524; DOI=10.1128/JB.186.9.2818-2828.2004;
RA  Haake D.A., Sutherland M.A., Kelley M.W., Dundoo M., Alt D.P.,
RA  Zuercher R.L.;
RT  "Molecular evolution and mosaicism of leptospiral outer membrane
RT  proteins involves horizontal DNA transfer.";
RL  J. Bacteriol. 186:2818-2828(2004).
DR  EMBL; AY461894; AAS21764.1; -.
FT  NON_TER      1      1
FT  NON_TER      259    259
SQ  SEQUENCE      259 AA; 28210 MW; A845A8035FE5A463 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 259;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      182 KPGQAP 187
      |||||
DB      63 KPGQAP 68

RESULT 15
Q7ARN6
ID  Q7ARN6      PRELIMINARY;      PRT;      259 AA.
AC  Q7ARN6;
DT  05-JUL-2004 (TReMBLrel. 27, Created)
DT  05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT  25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE  ORF1 (ORF 77) (ORF1; putative transposase).
OS  Yersinia pestis.
OG  Plasmid 9.5 kbp pesticin.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC  Enterobacteriaceae; Yersinia.
OX  NCBI_TaxID=632;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=6/69;
RA  Buchrieser C., Ruenloek C., Couve E., Frangeul L., Billault A.,
RA  Kunst F., Carniel E., Glaser P.;
RL  Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=EV76-6; PLASMID=9.5 kbp pesticin;
RX  MEDLINE=97221614; PubMed=9068660;
RT  McDonough K.A., Hare J.M.;
RT  "Homology with a repeated Yersinia pestis DNA sequence IS100
RT  correlates with pesticin sensitivity in Yersinia pseudotuberculosis.";
RL  J. Bacteriol. 179:2081-2085(1997).
DR  EMBL; AL031866; CAA21334.1; -.
DR  EMBL; AL031866; CAA21402.1; -.
DR  EMBL; U59875; AAC44982.1; -.
DR  GO; GO:0005524; F:ATP binding; IEA.
DR  GO; GO:000166; F:nucleotide binding; IEA.
DR  InterPro; IPR003593; AAA_ATPase.
DR  InterPro; IPR001270; Chaprin_c1pA/B.
DR  InterPro; IPR002611; IsetB_ATPbind.
DR  Pfam; PFO1695; IsetB.1.
DR  PRINTS; PR00300; CLPROTEASEA.
DR  SMART; SM00382; AAA.1.
KW  ATP-binding; Plasmid.
SQ  SEQUENCE      259 AA; 29295 MW; E72FD9033B80E379 CRC64;

```

```

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 259;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      186 APRLLI 191
      |||||
DB      164 APRLLI 169

```

Wed Sep 28 10:54:06 2005

us-09-936-702-3\_copy\_244\_502.oligsiz.rup

Page 7

Search completed: September 26, 2005, 13:14:50  
Job time : 170 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 10:28:40 ; Search time 0.001 Seconds  
(without alignments)  
548.562 Million cell updates/sec

Title: US-09-936-702-3  
Perfect score: 259  
Sequence: 1 QVQLTSGAEVKKPKGSSVKY.....RTFGQTRLEIKLVPRGSG 259

Scoring table: UNITARY AA  
Gapop 10.0 , Gapext 0.5

Searched: 27 seqs, 2118 residues

Total number of hits satisfying chosen parameters: 27

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 27 summaries

Database : rag.subdb.\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Description
Result No.	Score	Query Match Length	ID	
1	88	34.0	95 1 ABO27152	Human germline I1g
2	88	34.0	95 1 ABO27151	Human germline I1g
3	88	34.0	95 1 ADC99850	Germline VK gene L
4	88	34.0	95 1 ADC99846	Germline VK gene L
5	88	34.0	95 1 ADC99838	Germline VK gene L
6	88	34.0	95 1 ADD05454	Anti-MUC18 antibod
7	88	34.0	95 1 ADD05450	Anti-MUC18 antibod
8	88	34.0	95 1 ADP10095	VEGF antibody I1gh
9	88	34.0	95 1 ADP10197	Antibody light cha
10	88	34.0	95 1 ADP10198	Antibody light cha
11	88	34.0	95 1 ADP09990	Antibody light cha
12	88	34.0	95 1 ADP09991	Antibody light cha
13	88	34.0	95 1 ADP10096	VEGF antibody I1gh
14	88	34.0	95 1 ADP09880	Anti-MUC18 monoclo
15	88	34.0	95 1 ADP09888	Anti-MUC18 monoclo
16	88	34.0	95 1 ADP09892	Anti-MUC18 monoclo
17	88	34.0	95 1 ADJ80271	Vkappa gene locus
18	88	34.0	95 1 ADJ80272	Vkappa gene locus
19	88	34.0	95 1 ADO07341	Human antibody L16
20	88	34.0	95 1 ADO07340	Human antibody L2
21	30	11.6	32 1 ADR43440	PRL variant #11.
22	30	11.6	32 1 ADR31585	Murine TES-C21 ant
23	29	11.2	32 1 ADR43442	PRL variant #13.
24	29	11.2	32 1 ADR43441	PRL variant #12.
25	29	11.2	32 1 ADR31587	Murine TES-C21 ant
26	29	11.2	32 1 ADR31586	Murine TES-C21 ant
27	24	9.3	26 1 AAU90894	Insulin/insulin-I1

ALIGNMENTS

RESULT 1

ABO27152	ID	ABO27152 standard; protein; 95 AA.
XX	XX	
AC	XX	ABO27152;
XX	XX	
DT	XX	10-SEP-2003 (first entry)
XX	XX	
DE	XX	Human germline light chain variable region gene segment #32.
XX	XX	
KW	XX	Human: light chain variable region; VK: humanised antibody;
KW	XX	chimeric antibody; complementarity determining region; CDR;
XX	XX	canonical CDR structure type.
OS	XX	Homo sapiens.
PN	XX	US2003039649-A1.
PD	XX	27-FEB-2003.
PF	XX	12-JUL-2002; 2002US-00194975.
XX	XX	12-JUL-2001; 2001US-030511P.
PR	XX	(FOOT/) FOOT J.
PA	XX	
PI	XX	Foot J.
XX	XX	
DR	XX	WPI; 2003-492151/46.
XX	XX	
PT	XX	Making humanized antibody for converting antibody, by making chimeric
PT	XX	antibodies containing complementarity determining region from non-human
PT	XX	antibody and appropriate framework sequences of human antibodies.
PS	XX	Example 1; Fig 2; 31pp; English.
XX	XX	
CC	XX	The invention describes a method of making a humanised antibody,
CC	XX	comprising making chimeric antibodies containing a complementarity
CC	XX	determining region (CDR) from a non-human antibody and appropriate
CC	XX	framework sequences (I) of human antibodies. (I) is selected by using
CC	XX	canonical CDR structure types of non-human antibody in comparison to
CC	XX	germline canonical CDR structure types of human antibodies as the basis
CC	XX	for selection, for humanisation. The method is useful for making a
CC	XX	humanised antibody or a converted antibody. The method is applicable for
CC	XX	converting a subject antibody sequence of any subject species to a less
CC	XX	immunogenic form suitable for use in an object species. The method is
CC	XX	reliable for identifying suitable human framework sequences to support
CC	XX	non-human CDR regions and to provide humanised antibodies that retain
CC	XX	high antigen binding with low immunogenicity in humans, without the need
CC	XX	for direct comparison of framework sequences, without the need for
CC	XX	determining critically important amino acid residues in the framework,
CC	XX	and without the need for multiple iteration and construction to obtain
CC	XX	humanised antibodies with suitable therapeutic properties. The antibody
CC	XX	has high affinity and low immunogenicity without need for comparing
CC	XX	framework sequences between non-human and human antibodies. This sequence
CC	XX	represents a human light chain variable region gene segment used in the
CC	XX	creation of humanised antibodies
XX	XX	
SQ	XX	Sequence 95 AA;
XX	XX	
Query Match	34.0%;	Score 88; DB 1; Length 95;
Best Local Similarity	92.6%;	Pred. No. 6.6;
Matches	88;	Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY	387	ELELTSPATLSVSPGERATLSCRAAEISVSLDAWYQQRGAPRLTIYGASTRANGVA 446
DB	1	EIVMTSPPTLVSPGERATLSCRAQSVSNLAWYQQRGAPRLTIYGASTRANGVA 60
QY	447	RFGSGSGAEFTLTISLQSEDPFAVYVCOQYNNMP 481
DB	61	RFGSGSGTFTLTISLQSEDPFAVYVCOQYNNMP 95

RESULT 2

AB027151  
 ID ABO27151 standard; protein; 95 AA.  
 XX  
 AC ABO27151;  
 XX  
 DT 10-SEP-2003 (first entry)  
 DE Human germline light chain variable region gene segment #31.  
 XX  
 KM Human; light chain variable region; VK; humanised antibody;  
 KM chimeric antibody; complementarity determining region; CDR;  
 KM canonical CDR structure type.  
 XX  
 OS Homo sapiens.  
 PN US2003039649-A1.  
 PD 27-FEB-2003.  
 XX  
 PF 12-JUL-2002; 2002US-00194975.  
 XX  
 PR 12-JUL-2001; 2001US-0305111P.  
 XX  
 PA (FOOT/) FOOT/ J.  
 XX  
 PI Foote J;  
 XX  
 DR WPI; 2003-492151/46.  
 XX  
 PT Making humanized antibody for converting antibody, by making chimeric  
 PT antibodies containing complementarity determining region from non-human  
 PT antibody and appropriate framework sequences of human antibodies.  
 PS  
 XX Example 1; Fig 2; 31pp; English.

XX The invention describes a method of making a humanised antibody,  
 CC comprising making chimeric antibodies containing a complementarity  
 CC determining region (CDR) from a non-human antibody and appropriate  
 CC framework sequences (I) of human antibodies. (I) is selected by using  
 CC canonical CDR structure types of non-human antibody in comparison to  
 CC germline canonical CDR structure types of human antibodies as the basis  
 CC for selection, for humanisation. The method is useful for making a  
 CC humanised antibody or a converted antibody. The method is applicable for  
 CC converting a subject antibody sequence of any subject species to a less  
 CC immunogenic form suitable for use in an object species. The method is  
 CC reliable for identifying suitable human framework sequences to support  
 CC non-human CDR regions and to provide humanised antibodies that retain  
 CC high antigen binding with low immunogenicity in humans, without the need  
 CC for direct comparison of framework sequences, without the need for  
 CC determining critically important amino acid residues in the framework,  
 CC and without the need for multiple iteration and construction to obtain  
 CC humanised antibodies with suitable therapeutic properties. The antibody  
 CC has high affinity and low immunogenicity without need for comparing  
 CC framework sequences between non-human and human antibodies. This sequence  
 CC represents a human light chain variable region gene segment used in the  
 CC creation of humanised antibodies  
 XX  
 SQ Sequence 95 AA;

Query Match 34.0%; Score 88; DB 1; Length 95;  
 Best Local Similarity 92.6%; Pred. No. 6.6;  
 Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 387 ELLEIOTSPATLVSPEERATLSCRASESVSDDLAWYQKQKPGQAPRLIYGASTRATGVA 446  
 DB 1 EIVWTQSPATLVSPEERATLSCRASESVSDDLAWYQKQKPGQAPRLIYGASTRATGIPA 60  
 QY 447 RFGSGSGAEFTLTSSLSQSEDPFAVYCCQYNNMP 481  
 DB 61 RFGSGSGAEFTLTSSLSQSEDPFAVYCCQYNNMP 95

RESULT 3

ADC99850  
 ID ADC99850 standard; protein; 95 AA.  
 XX  
 AC ADC99850;  
 XX  
 DT 01-JAN-2004 (first entry)  
 DE Germline VK gene I2 region protein SEQ ID 79.  
 XX  
 KM anti-human MUC18 monoclonal antibody; heavy; light chain variable domain;  
 KM cytotoxic; melanoma; oesophageal; pancreatic; colorectal tumour;  
 KM cervical carcinoma; intraepithelial neoplasia; colorectal; breast;  
 KM lung cancer; germline VK region.  
 XX  
 OS Unidentified.  
 PN WO2003057838-A2.  
 PD 17-JUL-2003.  
 XX  
 PF 26-DEC-2002; 2002WO-US041581.  
 XX  
 PR 28-DEC-2001; 2001US-0346299P.  
 XX  
 PA (ABGE-) ABGENIX INC.  
 XX  
 PI Gudae J;  
 XX  
 DR WPI; 2003-587113/55.  
 XX  
 PT New human anti-MUC18 monoclonal antibodies, useful for treating a disease  
 PT or condition associated with expression of MUC18 in a patient, e.g.  
 PT tumors, cancers, and other malignancies.  
 PS  
 XX Example 2; SEQ ID NO 79; 78pp; English.

XX The invention relates to a novel isolated monoclonal antibody comprising  
 CC a heavy or light chain amino acid or a heavy or light chain variable  
 CC domain where the antibody binds to MUC18. The monoclonal antibody of the  
 CC invention demonstrates cytotoxic activity and may be useful for treating  
 CC a disease or condition associated with the expression of MUC18 on the  
 CC cell surface such as tumours, specifically melanoma, oesophageal,  
 CC pancreatic or colorectal tumours, carcinomas, particularly cervical  
 CC carcinomas and cervical intraepithelial neoplasia and cancers including  
 CC colorectal, breast or lung cancer, as well as other malignancies. The  
 CC current sequence is that of the germline VK gene region protein of the  
 CC invention used to analyse the anti-human MUC18 monoclonal antibody  
 CC sequences.  
 XX  
 SQ Sequence 95 AA;

Query Match 34.0%; Score 88; DB 1; Length 95;  
 Best Local Similarity 92.6%; Pred. No. 6.6;  
 Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 387 ELLEIOTSPATLVSPEERATLSCRASESVSDDLAWYQKQKPGQAPRLIYGASTRATGVA 446  
 DB 1 EIVWTQSPATLVSPEERATLSCRASESVSDDLAWYQKQKPGQAPRLIYGASTRATGIPA 60  
 QY 447 RFGSGSGAEFTLTSSLSQSEDPFAVYCCQYNNMP 481  
 DB 61 RFGSGSGAEFTLTSSLSQSEDPFAVYCCQYNNMP 95

RESULT 4  
 ID ADC99846 standard; protein; 95 AA.  
 XX  
 AC ADC99846;  
 XX  
 DT 01-JAN-2004 (first entry)  
 DE Germline VK gene I2 region protein SEQ ID 75.



KX		anti-human MUC18 monoclonal antibody; heavy; light chain variable domain;
KM	cystostatic; melanoma; oesophageal; pancreatic; colorectal tumour;	
KW	cervical carcinoma; intraepithelial neoplasia; colorectal; breast; lung cancer; germline VK region.	
OS	unidentified.	
XX		
PV	WO2003057938-A2.	
PD	17-JUL-2003.	
XX		
PE	26-DEC-2002; 2002MO-US041581.	
XX		
RH	28-DEC-2001; 2001US-0346299P.  (ABGE-) ABGENIX INC.	
PA	Gudas J;	
PI	WIPO; 2003-587113/55.	
DR	New human anti-MUC18 monoclonal antibodies, useful for treating a disease or condition associated with expression of MUC18 in a patient, e.g. tumors, cancers, and other malignancies.	
PT		
PT		
PS	Example 2; SEQ ID NO 75; 78pp; English.	
XX	The invention relates to a novel isolated monoclonal antibody comprising a heavy or light chain amino acid or a heavy or light chain variable domain where the antibody binds to MUC18. The monoclonal antibody of the invention demonstrates cytotoxic activity and may be useful for treating a disease or condition associated with the expression of MUC18 on the cell surface such as tumors, specifically melanoma, oesophageal, pancreatic or colorectal tumours, carcinomas, particularly cervical carcinomas and cervical intraepithelial neoplasia and cancers including colorectal, breast or lung cancer, as well as other malignancies. The current sequence is that of the germline VK gene region protein of the invention used to analyse the anti-human MUC18 monoclonal antibody sequences.	
SQ	Sequence 95 AA:	
Query Match	34.0%; Score 88; DB I; Length 95;	
Best Local Similarity	92.6%; Pred. No. 6.6;	
Matches	88; Conservative 0; Mismatches 7; Indels 0; Gaps 0,	
OY	387 EELTQSATLTSVSGERATLTSCRASSEYSDDLAWYOORPGAPRLLIYGASTRATGIPA 446 + + + + + DB 1 EIVMTQSPATLSVSGERATLTSCRASSQSVSNIMAYOORPGAPRLLIYGASTRATGIPA 60 + + + + +	
OY	447 RFSGSAGAEFTLLTTISLOSIEDFAFYIYYCOQQYNMP 481 + + + + + DB 61 RFSGSAGSTERTLLTISSLQSEDFAFYIYYCOQQINNP 95 + + + + +	
ID	ADCC99838 standard; protein; 95 AA.	
AC	ADCC99838;	
DT	01-JAN-2004 (first entry)	
DE	Germline VK gene L2 region protein SEQ ID 67.	
XX		
KW	anti-human MUC18 monoclonal antibody; heavy; light chain variable domain;	
KW	cystostatic; melanoma; oesophageal; pancreatic; colorectal tumour;	
KW	cervical carcinoma; intraepithelial neoplasia; colorectal; breast; lung cancer; germline VK region.	
OS	Unidentified.	
XX		

PN	MO2003057838-A2.
XX	
PD	17-JUL-2003.
XX	
PF	26-DEC-2002; 2002WO-USO41581.
XX	
PR	28-DEC-2001; 2001US-0346299P.
XX	
PA	(ABGE-) ABGENIX INC.
PI	Gudas J;
DR	WPI, 2003-587113/55.
XX	
PT	New human anti-MUC18 monoclonal antibodies, useful for treating a disease or condition associated with expression of MUC18 in a patient, e.g. tumors, cancers, and other malignancies.
PS	Example 2; SEQ ID NO 67; 78pp; English.
XX	
CC	The invention relates to a novel isolated monoclonal antibody comprising a heavy or light chain amino acid or a heavy or light chain variable domain where the antibody binds to MUC18. The monoclonal antibody of the invention demonstrates cytostatic activity and may be useful for treating a disease or condition associated with the expression of MUC18 on the cell surface such as tumours, specifically melanoma, oesophageal, pancreatic or colorectal tumours, carcinomas, particularly cervical carcinoma and cervical intraepithelial neoplasia and cancers including colorectal, breast or lung cancer, as well as other malignancies. The current sequence is that of the germline VK gene region protein of the invention used to analyse the anti-human MUC18 monoclonal antibody sequences.
SQ	Sequence 95 AA;
Query Match	34.0%; Score 88; DB 1; Length 95;
Best Local Similarity	92.6%; Pred. NO. 6.6;
Matches	88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY	387 ELELTQSPATLVSPPERATLSCRASESVSSDLAWYOQKPGQAPRLLIYGASTRATGVPA 446           1 EIWMTQSATISVSPGERATISCRASQSVSSNLAWYQCKPGQARLLIYGASTRATGIPA 60
OY	447 RFGSGSGAEFTLTITSSIQSEDPFVVYYCOQYNMP 481           61 RFGSGSGCTEFTLTITSSIQSEDPFVVYYCOQYNMP 95
DB	
RESULT 6	
ADD05454	ADD05454 standard; protein; 95 AA.
AC	ADD05454;
DT	01-JAN-2004 (first entry)
DE	Anti-MUC18 antibody light chain variable region I2 protein, SEQ ID 79.
KM	monoclonal antibody; tumour; MUC18; proliferation; cytotaxtic; vaccine; antigen; tumour metastasis; melanoma; metastatic; human; light chain.
OS	Unidentified.
PN	MO2003057006-A2.
PD	17-JUL-2003.
PF	26-DEC-2002; 2002WO-USO41582.
PR	28-DEC-2001; 2001US-0346460P.
PA	(ABGE-) ABGENIX INC.
XX	

PI Gudaa J, Bar-Eli M;  
 XX WPI; 2003-577496/54.  
 XX  
 PT Use of monoclonal antibodies against MUC18 antigen, for diagnosing and  
 PT treating tumors, inhibiting tumor growth, inhibiting cell invasion  
 PT associated with melanoma, or increasing survival of an animal having a  
 PT metastatic tumor.  
 PS Disclosure; SEQ ID NO 79; 87bp; English.  
 XX  
 CC The invention relates to a novel monoclonal antibody used for inhibiting  
 CC tumor growth in an animal. The tumor inhibition process comprises  
 CC selecting an animal in need of treatment for a tumor, providing a  
 CC monoclonal antibody comprising a heavy chain amino acid, where the  
 CC antibody consists of any one of 10 fully defined sequences of 117-123  
 CC amino acids given in the specification, and where the monoclonal antibody  
 CC binds MUC18, and contacting the tumor with the antibody resulting in  
 CC inhibited proliferation of the cells. The monoclonal antibody has  
 CC cytostatic and can be used in the production of a vaccine. The monoclonal  
 CC antibodies against the MUC18 antigen are useful for diagnosing and  
 CC treating tumors, inhibiting tumor growth (e.g. melanoma, lung tumor or  
 CC tumor metastasis), inhibiting cell invasion associated with melanoma, or  
 CC increasing survival of an animal having a metastatic tumor. This  
 CC sequence represents an anti-MUC18 antibody light chain, variable region,  
 CC protein of the invention.  
 XX  
 SQ Sequence 95 AA;  
 Query Match 34.0%; Score 88; DB 1; Length 95;  
 Best Local Similarity 92.6%; Pred. No. 6.6;  
 Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 387 ELELTSPATLSVSPGERATLSGRASESVSSDLAWYQKPGQAPRLIYGASTRATGVA 446  
 Db 1 EIVMTSPATLSVSPGERATLSGRASESVSSDLAWYQKPGQAPRLIYGASTRATGVA 60  
 QY 447 RFGSGSGAEFTLTITSSLSQSEDFAVYYCQQYNWP 481  
 Db 61 RFGSGSGTEFTLTITSSLSQSEDFAVYYCQQYNWP 95  
 Db  
 QY 447 RFGSGSGAEFTLTITSSLSQSEDFAVYYCQQYNWP 481  
 Db 61 RFGSGSGTEFTLTITSSLSQSEDFAVYYCQQYNWP 95  
 Db  
 RESULT 7  
 ADD05450  
 ID ADD05450 standard; protein; 95 AA.  
 XX  
 AC ADD05450;  
 XX  
 DT 01-JAN-2004 (first entry)  
 XX  
 DE Anti-MUC18 antibody light chain variable region L2 protein, SEQ ID 75.  
 XX  
 KM monoclonal antibody; tumour; MUC18; proliferation; cytostatic; vaccine;  
 KM antigen; tumour metastasis; melanoma; metastatic; human; light chain.  
 XX  
 OS Unidentified.  
 OS  
 PN WO2003057006-A2.  
 XX  
 PD 17-JUL-2003.  
 XX  
 PF 26-DEC-2002; 2002WO-US041582.  
 XX  
 PR 28-DEC-2001; 2001US-0346460P.  
 XX  
 PA (ABGE-) ABGENIX INC.  
 XX  
 PI Gudaa J, Bar-Eli M;  
 XX  
 DR WPI; 2003-577496/54.  
 XX  
 PT Use of monoclonal antibodies against MUC18 antigen, for diagnosing and  
 PT treating tumors, inhibiting tumor growth, inhibiting cell invasion

PT associated with melanoma, or increasing survival of an animal having a  
 PT metastatic tumor.  
 XX  
 PS Disclosure; SEQ ID NO 75; 87bp; English.  
 XX  
 CC The invention relates to a novel monoclonal antibody used for inhibiting  
 CC tumor growth in an animal. The tumor inhibition process comprises  
 CC selecting an animal in need of treatment for a tumor, providing a  
 CC monoclonal antibody comprising a heavy chain amino acid, where the  
 CC antibody consists of any one of 10 fully defined sequences of 117-123  
 CC amino acids given in the specification, and where the monoclonal antibody  
 CC binds MUC18, and contacting the tumor with the antibody resulting in  
 CC inhibited proliferation of the cells. The monoclonal antibody has  
 CC cytostatic and can be used in the production of a vaccine. The monoclonal  
 CC antibodies against the MUC18 antigen are useful for diagnosing and  
 CC treating tumors, inhibiting tumor growth (e.g. melanoma, lung tumor or  
 CC tumor metastasis), inhibiting cell invasion associated with melanoma, or  
 CC increasing survival of an animal having a metastatic tumor. This  
 CC sequence represents an anti-MUC18 antibody light chain, variable region,  
 CC protein of the invention.  
 XX  
 SQ Sequence 95 AA;  
 Query Match 34.0%; Score 88; DB 1; Length 95;  
 Best Local Similarity 92.6%; Pred. No. 6.6;  
 Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 387 ELELTSPATLSVSPGERATLSGRASESVSSDLAWYQKPGQAPRLIYGASTRATGVA 446  
 Db 1 EIVMTSPATLSVSPGERATLSGRASESVSSDLAWYQKPGQAPRLIYGASTRATGVA 60  
 QY 447 RFGSGSGAEFTLTITSSLSQSEDFAVYYCQQYNWP 481  
 Db 61 RFGSGSGTEFTLTITSSLSQSEDFAVYYCQQYNWP 95  
 Db  
 QY 447 RFGSGSGAEFTLTITSSLSQSEDFAVYYCQQYNWP 481  
 Db 61 RFGSGSGTEFTLTITSSLSQSEDFAVYYCQQYNWP 95  
 Db  
 RESULT 8  
 ADF10095  
 ID ADF10095 standard; protein; 95 AA.  
 XX  
 AC ADF10095;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE VEGF antibody light chain variable region VLK\_3-15.  
 XX  
 KM Antibody; stability; solubility; antigen binding affinity;  
 KM variable region; human; VEGF.  
 XX  
 OS Homo sapiens.  
 OS  
 PN WO2003074679-A2.  
 XX  
 PD 12-SEP-2003.  
 XX  
 PF 03-MAR-2003; 2003WO-US006598.  
 XX  
 PR 01-MAR-2002; 2002US-0360843P.  
 PR 29-MAY-2002; 2002US-0384197P.  
 XX  
 PA (XENC-) XENCOR.  
 XX  
 PI Lazar GA, Desjarlais JR, Marshall SA, Dahiya B;  
 XX  
 DR WPI; 2003-722066/68.  
 XX  
 PT Computer optimization of physicochemical properties of antibodies  
 PT comprises analyzing the interactions of amino acids at variable  
 PT positions.  
 XX  
 PS Example 6; Fig 16d; 135bp; English.  
 XX  
 PT The present invention relates to a method for optimizing at least one

CC physico-chemical property of an antibody by a computational screening  
 CC method. The method comprises: receiving a template antibody structure;  
 CC selecting at least one variable position belonging to the antibody  
 CC structure; selecting at least one amino acid to be considered at the  
 CC variable position(s); analyzing the interaction of each selected amino  
 CC acid at each variable position with at least part of the remainder of the  
 CC antibody, including the selected amino acids at other variable positions;  
 CC and identifying a set of at least one antibody sequence with at least one  
 CC optimized physico-chemical property. The method is useful for optimizing  
 CC the physico-chemical properties of an antibody, especially the stability,  
 CC solubility, or antigen binding affinity. The optimized antibody may be  
 CC useful for treating a patient. The present sequence is an antibody  
 CC variable region sequence used to illustrate the invention.

CC Sequence 95 AA:

Query Match 34.0%; Score 88; DB 1; Length 95;  
 Best Local Similarity 92.6%; Pred. No. 6.6;  
 Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELELTSPATLSVSPGERATLSCRASSEVSDDLAWYQKPGQAPRLIYGASTRATGVA 446  
 DB 1 EIVMTQSPATLSVSPGERATLSCRASQSVSNLAWYQKPGQAPRLIYGASTRATGIPA 60

QY 447 RFGSGSGAEFTLTISLQSEDFAVYCCQYNNWP 481  
 DB 61 RFGSGSGTEFTLTISLQSEDFAVYCCQYNNWP 95

RESULT 9  
 ID ADF10197 standard; protein; 95 AA.

XX ADF10197;

XX 12-FEB-2004 (first entry)

XX Antibody light chain variable region VLK\_3-15.

XX Antibody: stability; solubility; antigen binding affinity;

XX variable region; human.

XX Homo sapiens.

XX MO2003074679-A2.

XX 12-SEP-2003.

XX 03-MAR-2003; 2003MO-US006598.

XX 01-MAR-2002; 2002US-0360843P.

XX 29-MAY-2002; 2002US-0384197P.

XX (XENC-) XENCOR.

PI Lazar GA, Desjarlais JR, Marshall SA, Dahiyat B;

XX WPI; 2003-722066/68.

PT Computer optimization of physicochemical properties of antibodies  
 PT comprises analyzing the interactions of amino acids at variable  
 PT positions.

XX Example 16; Fig 40b; 135pp; English.

XX The present invention relates to a method for optimizing at least one  
 CC physico-chemical property of an antibody by a computational screening  
 CC method. The method comprises: receiving a template antibody structure;  
 CC selecting at least one variable position belonging to the antibody  
 CC structure; selecting at least one amino acid to be considered at the  
 CC variable position(s); analyzing the interaction of each selected amino  
 CC acid at each variable position with at least part of the remainder of the  
 CC antibody, including the selected amino acids at other variable positions;

CC and identifying a set of at least one antibody sequence with at least one  
 CC optimized physico-chemical property. The method is useful for optimizing  
 CC the physico-chemical properties of an antibody, especially the stability,  
 CC solubility, or antigen binding affinity. The optimized antibody may be  
 CC useful for treating a patient. The present sequence is an antibody  
 CC variable region sequence used to illustrate the invention.

CC Sequence 95 AA:

Query Match 34.0%; Score 88; DB 1; Length 95;  
 Best Local Similarity 92.6%; Pred. No. 6.6;  
 Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELELTSPATLSVSPGERATLSCRASSEVSDDLAWYQKPGQAPRLIYGASTRATGVA 446  
 DB 1 EIVMTQSPATLSVSPGERATLSCRASQSVSNLAWYQKPGQAPRLIYGASTRATGIPA 60

QY 447 RFGSGSGAEFTLTISLQSEDFAVYCCQYNNWP 481  
 DB 61 RFGSGSGTEFTLTISLQSEDFAVYCCQYNNWP 95

RESULT 10  
 ID ADF10198 standard; protein; 95 AA.

XX ADF10198;

XX 12-FEB-2004 (first entry)

XX Antibody light chain variable region VLK\_3D-15.

XX Antibody: stability; solubility; antigen binding affinity;

XX variable region; human.

XX Homo sapiens.

XX MO2003074679-A2.

XX 12-SEP-2003.

XX 03-MAR-2003; 2003MO-US006598.

XX 01-MAR-2002; 2002US-0360843P.

XX 29-MAY-2002; 2002US-0384197P.

XX (XENC-) XENCOR.

PI Lazar GA, Desjarlais JR, Marshall SA, Dahiyat B;

XX WPI; 2003-722066/68.

PT Computer optimization of physicochemical properties of antibodies  
 PT comprises analyzing the interactions of amino acids at variable  
 PT positions.

XX Example 16; Fig 40b; 135pp; English.

XX The present invention relates to a method for optimizing at least one  
 CC physico-chemical property of an antibody by a computational screening  
 CC method. The method comprises: receiving a template antibody structure;  
 CC selecting at least one variable position belonging to the antibody  
 CC structure; selecting at least one amino acid to be considered at the  
 CC variable position(s); analyzing the interaction of each selected amino  
 CC acid at each variable position with at least part of the remainder of the  
 CC antibody, including the selected amino acids at other variable positions;  
 CC and identifying a set of at least one antibody sequence with at least one  
 CC optimized physico-chemical property. The method is useful for optimizing  
 CC the physico-chemical properties of an antibody, especially the stability,  
 CC solubility, or antigen binding affinity. The optimized antibody may be  
 CC useful for treating a patient. The present sequence is an antibody  
 CC variable region sequence used to illustrate the invention.

Seq	Sequence	95 AA;
Query	Match	34.0%; Score 88; DB 1; Length 95;
Best Local	Similarity	92.6%; Pred. No. 6.6;
Matches	88; Conservative	0; Mismatches 7; Indels 0; Gaps 0;
Qy	387	ELELTQSPATISVSGGERATISCRASEVSSDLAWYQOKPGQAPRLIIYGASTRATGVPA 446
Db	1	EIVMTQSPATISVSGGERATISCRASGVSSNLAWYQOKPGQAPRLIIYGASTRATGIPA 60
Qy	447	RFGSGSGGAFFLTITISLQSEDFAVYYCOQYNWNP 481
Db	61	RFGSGSGCTEFTLTITISLQSEDFAVYYCOQYNWNP 95
RESULT 11		
ID	ADF09990	standard; protein; 95 AA.
XX	ADF09990;	
AC	ADF09990;	
XX		
DT	12-FEB-2004	(first entry)
XX		
DE	Antibody light chain variable region VLK_3-15.	
XX		
KW	Antibody; stability; solubility; antigen binding affinity;	
KW	variable region; human.	
XX		
OS	Homo sapiens.	
XX		
PN	WO2003074679-A2.	
PD	12-SEP-2003.	
XX		
PF	03-MAR-2003; 2003WO-US006598.	
PR	01-MAR-2002; 2002US-0360843P.	
XX	29-MAY-2002; 2002US-0384197P.	
XX		
PA	(XENC-) XENCOR.	
XX		
PI	Lazar GA, Desjarlais JR, Marshall SA, Dahiyat B;	
DR	WPI; 2003-722066/68.	
XX		
PT	Computer optimization of physicochemical properties of antibodies	
PT	comprises analyzing the interactions of amino acids at variable	
PT	positions.	
XX		
PS	Disclosure; Fig 2b; 135p; English.	
XX		
CC	The present invention relates to a method for optimizing at least one	
CC	physico-chemical property of an antibody by a computational screening	
CC	method. The method comprises: receiving a template antibody structure;	
CC	selecting at least one variable position belonging to the antibody	
CC	structure; selecting at least one amino acid to be considered at the	
CC	variable position(s); analyzing the interaction of each selected amino	
CC	acid at each variable position with at least part of the remainder of the	
CC	antibody, including the selected amino acids at other variable positions;	
CC	and identifying a set of at least one antibody sequence with at least one	
CC	optimized physico-chemical property. The method is useful for optimizing	
CC	the physico-chemical properties of an antibody, especially the stability,	
CC	solubility, or antigen binding affinity. The optimized antibody may be	
CC	useful for treating a patient. The present sequence is an antibody	
CC	variable region sequence used to illustrate the invention.	
XX		
SQ	Sequence 95 AA;	
Query	Match	34.0%; Score 88; DB 1; Length 95;
Best Local	Similarity	92.6%; Pred. No. 6.6;
Matches	88; Conservative	0; Mismatches 7; Indels 0; Gaps 0;
Qy	387	ELELTQSPATISVSGGERATISCRASEVSSDLAWYQOKPGQAPRLIIYGASTRATGVPA 446

[illegible]

RESULT 13  
ADFI0096  
XX ADFI0096 standard; protein; 95 AA.  
XX AC ADFI0096;  
XX DT 12-FEB-2004 (first entry)  
XX DE VEGF antibody light chain variable region VLK\_3D-15.  
XX DE VEGF antibody light chain variable region VLK\_3D-15.  
XX KW Antibody; stability; solubility; antigen binding affinity;  
XX KW variable region; human; VEGF.  
XX OS Homo sapiens.  
XX PN WO2003074679-A2.  
XX PD 12-SEP-2003.  
XX PF 03-MAR-2003; 2003WO-US006598.  
XX PR 01-MAR-2002; 2002US-0360843P.  
XX PR 29-MAY-2002; 2002US-0384197P.  
XX PA (XENC-) XENCOR.  
XX PI Lazar GA, Desjarlais JR, Marshall SA, Dahiyat B;  
XX WPI; 2003-722066/68.  
XX DR WPI; 2003-722066/68.  
XX PT Computer optimization of physicochemical properties of antibodies  
XX PT comprises analyzing the interactions of amino acids at variable  
XX PT positions.  
XX PS Example 6; Fig 16b; 135pp; English.  
XX PS The present invention relates to a method for optimizing at least one  
XX CC physico-chemical property of an antibody by a computational screening  
XX CC method. The method comprises: receiving a template antibody structure;  
XX CC selecting at least one variable position belonging to the antibody  
XX CC structure; selecting at least one amino acid to be considered at the  
XX CC variable position(s); analyzing the interaction of each selected amino  
XX CC acid at each variable position with at least part of the remainder of the  
XX CC antibody, including the selected amino acids at other variable positions;  
XX CC and identifying a set of at least one antibody sequence with at least one  
XX CC optimized physico-chemical property. The method is useful for optimizing  
XX CC the physico-chemical properties of an antibody, especially the stability,  
XX CC solubility, or antigen binding affinity. The optimized antibody may be  
XX CC useful for treating a patient. The present sequence is an antibody  
XX CC variable region sequence used to illustrate the invention.  
XX SQ Sequence 95 AA;  
SQ Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 6.6;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 387 ELELTGSPATLSVSPGERATLSCRASVSVDLAWYQQRGQAPRLIIYGASTRATGVP 446  
DB 1 EIVMTGSPATLSVSPGERATLSCRASVSVDLAWYQQRGQAPRLIIYGASTRATGIP 60  
QY 447 RFSGSGSGAEFTLTISLSQSEDPFAYVYCCQYNNMP 481  
DB 61 RFSGSGSGAEFTLTISLSQSEDPFAYVYCCQYNNMP 95  
RESULT 14  
ADP09880  
XX ADP09880 standard; protein; 95 AA.  
XX AC ADF09880;  
XX XX

DT 12-FEB-2004 (first entry)  
XX DE Anti-MUC18 monoclonal antibody-related protein #14.  
XX DE Anti-MUC18 monoclonal antibody-related protein #14.  
XX KW cell proliferation inhibition; MUC18 tumour antigen;  
XX KW anti-MUC18 monoclonal antibody; tumour metastasis inhibition; tumour;  
XX KW carcinoma; cancer; malignancy.  
XX OS Unidentified.  
XX PN WO2003057837-A2.  
XX PD 17-JUL-2003.  
XX PF 26-DEC-2002; 2002WO-US041580.  
XX PR 28-DEC-2001; 2001US-0346414P.  
XX PA (ABGE-) ABGENIX INC.  
XX PI Gudas J;  
XX WPI; 2003-598367/56.  
XX DR WPI; 2003-598367/56.  
XX PT Inhibiting cell proliferation associated with expression of MUC18 tumor  
XX PT antigen, involves incubating and inhibiting cell by administering anti-  
XX PT MUC18 monoclonal antibody.  
XX PS Example 2; SEQ ID NO 67; 83pp; English.  
XX PS The invention comprises a method for inhibiting cell proliferation  
XX CC associated with expression of MUC18 tumour antigen. The method involves  
XX CC administering anti-MUC18 monoclonal antibody. The method of the invention  
XX CC is useful for inhibiting cell (e.g. melanoma or tumour cell)  
XX CC proliferation associated with the expression of MUC18 tumour antigen, the  
XX CC method is preferably useful for inhibiting tumour metastasis. The method  
XX CC is useful for inhibiting cell proliferation in patients with tumours,  
XX CC carcinomas, cancer and other malignancies. The present amino acid  
XX CC sequence is used in an alignment with an MUC18 tumour antigen-specific  
XX CC monoclonal antibody of the invention.  
XX SQ Sequence 95 AA;  
SQ Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 6.6;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 387 ELELTGSPATLSVSPGERATLSCRASVSVDLAWYQQRGQAPRLIIYGASTRATGVP 446  
DB 1 EIVMTGSPATLSVSPGERATLSCRASVSVDLAWYQQRGQAPRLIIYGASTRATGIP 60  
QY 447 RFSGSGSGAEFTLTISLSQSEDPFAYVYCCQYNNMP 481  
DB 61 RFSGSGSGAEFTLTISLSQSEDPFAYVYCCQYNNMP 95  
RESULT 15  
ADP09888  
XX ADP09888 standard; protein; 95 AA.  
XX AC ADF09888;  
XX DT 12-FEB-2004 (first entry)  
XX DE Anti-MUC18 monoclonal antibody-related protein #18.  
XX DE Anti-MUC18 monoclonal antibody-related protein #18.  
XX KW cell proliferation inhibition; MUC18 tumour antigen;  
XX KW anti-MUC18 monoclonal antibody; tumour metastasis inhibition; tumour;  
XX KW carcinoma; cancer; malignancy.  
XX OS Unidentified.  
XX PN WO2003057837-A2.  
XX XX

```

XX 17-JUL-2003.
PD
PT 26-DEC-2002; 2002WO-US041580.
PF
PR 28-DEC-2001; 2001US-0346414P.
XX
XX (ABGE-) ABGENIX INC.
PA
PI Gudae J;
XX
XX WPI; 2003-598367/56.
DR
XX Inhibiting cell proliferation associated with expression of MUC18 tumor
PT antigen, involves incubating and inhibiting cell by administering anti-
PT MUC18 monoclonal antibody.
XX
XX Example 2; SEQ ID NO 75; 83pp; English.
PS
XX The invention comprises a method for inhibiting cell proliferation
CC associated with expression of MUC18 tumour antigen. The method involves
CC administering anti-MUC18 monoclonal antibody. The method of the invention
CC is useful for inhibiting cell (e.g. melanoma or tumour cell)
CC proliferation associated with the expression of MUC18 tumour antigen, the
CC method is preferably useful for inhibiting tumour metastasis. The method
CC is useful for inhibiting cell proliferation in patients with tumours,
CC carcinomas, cancer and other malignancies. The present amino acid
CC sequence is used in an alignment with an MUC18 tumour antigen-specific
CC monoclonal antibody of the invention.
XX
XX Sequence 95 AA:
SQ
XX
XX Query Match 34.0%; Score 88; DB 1; Length 95;
XX Best Local Similarity 92.6%; Pred. No. 6.6;
XX Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 387 ELELTGSPATLVSFGERATLSCRASBSVSDLAWYQKPGQAPRLIYGASTRATGVA 446
DB 1 EIVMTGSPATLVSFGERATLSCRASQSVSNLAWYQKPGQAPRLIYGASTRATGIPA 60
QY 447 RFGSGSGAEFTLTISLQSEDPFAYVYCCQYNNWP 481
DB 61 RFGSGSGTEFTLTISLQSEDPFAYVYCCQYNNWP 95
XX
XX RESULT 16
XX ADF09892
ID ADF09892 standard; protein; 95 AA.
XX
XX ADF09892;
AC
XX
XX 12-FEB-2004 (first entry)
DT
XX
XX Anti-MUC18 monoclonal antibody-related protein #20.
DE
XX
XX cell proliferation inhibition; MUC18 tumour antigen;
KM anti-MUC18 monoclonal antibody; tumour metastasis inhibition; tumour;
KM carcinoma; cancer; malignancy.
XX
XX Unidentified.
OS
XX
XX WO2003057837-A2.
PN
XX
XX 17-JUL-2003.
PD
XX
XX 26-DEC-2002; 2002WO-US041580.
PF
XX
XX 28-DEC-2001; 2001US-0346414P.
PR
XX
XX (ABGE-) ABGENIX INC.
PA
XX
XX Gudae J;
PI
XX

```

```

DR WPI; 2003-598367/56.
XX
XX Inhibiting cell proliferation associated with expression of MUC18 tumor
PT antigen, involves incubating and inhibiting cell by administering anti-
PT MUC18 monoclonal antibody.
XX
XX Example 2; SEQ ID NO 79; 83pp; English.
PS
XX The invention comprises a method for inhibiting cell proliferation
CC associated with expression of MUC18 tumour antigen. The method involves
CC administering anti-MUC18 monoclonal antibody. The method of the invention
CC is useful for inhibiting cell (e.g. melanoma or tumour cell)
CC proliferation associated with the expression of MUC18 tumour antigen, the
CC method is preferably useful for inhibiting tumour metastasis. The method
CC is useful for inhibiting cell proliferation in patients with tumours,
CC carcinomas, cancer and other malignancies. The present amino acid
CC sequence is used in an alignment with an MUC18 tumour antigen-specific
CC monoclonal antibody of the invention.
XX
XX Sequence 95 AA:
SQ
XX
XX Query Match 34.0%; Score 88; DB 1; Length 95;
XX Best Local Similarity 92.6%; Pred. No. 6.6;
XX Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 387 ELELTGSPATLVSFGERATLSCRASBSVSDLAWYQKPGQAPRLIYGASTRATGVA 446
DB 1 EIVMTGSPATLVSFGERATLSCRASQSVSNLAWYQKPGQAPRLIYGASTRATGIPA 60
QY 447 RFGSGSGAEFTLTISLQSEDPFAYVYCCQYNNWP 481
DB 61 RFGSGSGTEFTLTISLQSEDPFAYVYCCQYNNWP 95
XX
XX RESULT 17
XX ADJ80271
ID ADJ80271 standard; protein; 95 AA.
XX
XX ADJ80271;
AC
XX
XX 06-MAY-2004 (first entry)
DT
XX
XX V kappa gene locus amino acid sequence #31.
DE
XX
XX hybrid antibody; antibody; framework region; homology; immunogenicity.
KM
XX
XX Homo sapiens.
OS
XX
XX WO2003048321-A2.
PN
XX
XX 12-JUN-2003.
PD
XX
XX 03-DEC-2002; 2002WO-US038450.
PF
XX
XX 03-DEC-2001; 2001US-0336591P.
PR
XX
XX (ALEX-) ALEXION PHARM INC.
PA
XX
XX Rother R, Wu D;
PI
XX
XX WPI; 2003-513753/48.
DR
XX
XX Producing a hybrid antibody or hybrid antibody fragment by operatively
PT linking the selected framework sequences to one or more complementarity
PT determining regions of the initial antibody.
XX
XX Disclosure; SEQ ID NO 31; 77pp; English.
PS
XX
XX The invention relates to a method of producing a hybrid antibody or
CC hybrid antibody fragment by: (i) providing an initial antibody having
CC specificity for a target; (ii) determining the sequence of a variable
CC region of the initial antibody; (iii) selecting a first component of the
CC variable region consisting of FR1, FR2, FR3 and FR4; (iv) comparing the

```

CC sequence of the first component to sequences contained in a reference  
 CC database of antibody sequences or antibody fragment sequences from a  
 CC target species; (v) selecting a sequence from an antibody in the database  
 CC which demonstrates a high degree of homology to the first component; (vi)  
 CC selecting a second component of the variable region which is different  
 CC than the first component, the second component selected from the group  
 CC consisting of FR1, FR2, FR3 and FR4; (vii) comparing the sequence of the  
 CC second component to sequences contained in a reference database of  
 CC antibody sequences or antibody fragment sequences from the target species  
 CC ; (viii) selecting a sequence from the database which demonstrates a high  
 CC degree of homology to the second component and which is from a different  
 CC antibody than the selected antibody; and (ix) operatively linking the  
 CC selected framework sequences to one or more complementarily determining  
 CC regions (CDRs) of the initial antibody to produce a hybrid antibody or  
 CC hybrid antibody fragment. The method is useful for producing a hybrid  
 CC antibody or hybrid antibody fragment (claimed). The antibody and  
 CC fragments are useful for therapeutic and diagnostic purposes. The method  
 CC uses entire framework regions from a single antibody variable heavy or  
 CC variable light chain to receive the CDRs. This produces antibodies that  
 CC are highly homologous and exhibit reduced immunogenicity while  
 CC maintaining an optimum binding profile. This sequence represents the  
 CC amino acid sequence of an antibody from the V kappa gene locus.

XX Sequence 95 AA;

Query Match 34.0%; Score 88; DB 1; Length 95;

Best Local Similarity 92.6%; Pred. No. 6.6;

Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 387 ELELTPSPATLSVSPGERATLSCRASBSVSSDLAWYQKQKQAPRLIYGASTRATGVA 446

Db 1 EIVWTGSPATLSVSPGERATLSCRASQSVSSNLAWYQKQKQAPRLIYGASTRATGIPA 60

Qy 447 RFGSGSGAEFTLTISLSQSEDFAVYYCCQYNMNP 481

Db 61 RFGSGSGTEFTLTISLSQSEDFAVYYCCQYNMNP 95

RESULT 18

ADJ80272 standard; protein; 95 AA.

XX AC ADJ80272;

DT 06-MAY-2004 (first entry)

DE V kappa gene locus antibody amino acid sequence #32.

XX hybrid antibody; antibody; framework region; homology; immunogenicity.

OS Homo sapiens.

XX WO2003048321-A2.

PD 12-JUN-2003.

PF 03-DEC-2002; 2002MO-US038450.

PR 03-DEC-2001; 2001US-0336591P.

PA (ALEX-) ALEXION PHARM INC.

PI Rother R, Wu D;

DR WPI; 2003-513753/48.

XX Producing a hybrid antibody or hybrid antibody fragment by operatively  
 PT linking the selected framework sequences to one or more complementarily  
 PT determining regions of the initial antibody.

XX Disclosure; SEQ ID NO 32; 77pp; English.

XX The invention relates to a method of producing a hybrid antibody or

CC hybrid antibody fragment by: (i) providing an initial antibody having  
 CC specificity for a target; (ii) determining the sequence of a variable  
 CC region of the initial antibody; (iii) selecting a first component of the  
 CC variable region consisting of FR1, FR2, FR3 and FR4; (iv) comparing the  
 CC sequence of the first component to sequences contained in a reference  
 CC database of antibody sequences or antibody fragment sequences from a  
 CC target species; (v) selecting a sequence from an antibody in the database  
 CC which demonstrates a high degree of homology to the first component; (vi)  
 CC selecting a second component of the variable region which is different  
 CC than the first component, the second component selected from the group  
 CC consisting of FR1, FR2, FR3 and FR4; (vii) comparing the sequence of the  
 CC second component to sequences contained in a reference database of  
 CC antibody sequences or antibody fragment sequences from the target species  
 CC ; (viii) selecting a sequence from the database which demonstrates a high  
 CC degree of homology to the second component and which is from a different  
 CC antibody than the selected antibody; and (ix) operatively linking the  
 CC selected framework sequences to one or more complementarily determining  
 CC regions (CDRs) of the initial antibody to produce a hybrid antibody or  
 CC hybrid antibody fragment. The method is useful for producing a hybrid  
 CC antibody or hybrid antibody fragment (claimed). The antibody and  
 CC fragments are useful for therapeutic and diagnostic purposes. The method  
 CC uses entire framework regions from a single antibody variable heavy or  
 CC variable light chain to receive the CDRs. This produces antibodies that  
 CC are highly homologous and exhibit reduced immunogenicity while  
 CC maintaining an optimum binding profile. This sequence represents the  
 CC amino acid sequence of an antibody from the V kappa gene locus.

XX Sequence 95 AA;

Query Match 34.0%; Score 88; DB 1; Length 95;

Best Local Similarity 92.6%; Pred. No. 6.6;

Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 387 ELELTPSPATLSVSPGERATLSCRASBSVSSDLAWYQKQKQAPRLIYGASTRATGVA 446

Db 1 EIVWTGSPATLSVSPGERATLSCRASQSVSSNLAWYQKQKQAPRLIYGASTRATGIPA 60

Qy 447 RFGSGSGAEFTLTISLSQSEDFAVYYCCQYNMNP 481

Db 61 RFGSGSGTEFTLTISLSQSEDFAVYYCCQYNMNP 95

RESULT 19

ADJ007341 standard; protein; 95 AA.

XX AC ADJ007341;

DT 15-JUL-2004 (first entry)

DE Human antibody L16 light chain variable region.

XX Catalytic antibody; human; antibody; gene therapy.

OS Homo sapiens.

XX Location/Qualifiers

FT 1..23 /label= FR1

FT 24..34 /note= "Framework region 1"

FT /label= CDR1 /note= "Complementarity determining region 1"

FT 35..49 /label= FR2

FT 50..56 /note= "Framework region 2"

FT /label= CDR2 /note= "Complementarity determining region 2"

FT 57..68 /label= FR3

FT /note= "Framework region 3"

```

FT      /label= CDR3
XT      /note= "Complementarity determining region 3"
XX
XX      MO2004033658-A2.
XX
XX      22-APR-2004.
XX
XX      09-OCT-2003; 2003WO-US032214.
XX
XX      10-OCT-2002; 2002US-0417979P.
XX
XX      (INTE-) INTEGRIGEN INC.
XX
XX      Smider V, Larrick JW;
XX
XX      WPI; 2004-340921/31.
XX
XX      Novel recombinant catalytic polypeptide useful for cleaving target
PT      proteins or for treating or preventing cancers, comprises a human
PT      antibody light chain operably joined to a heterologous antibody heavy
PT      chain.
XX
XX      Disclosure; Fig 3; 92pp; English.
XX
XX      The present sequence is the light chain variable region of human antibody
XX      116. This is one of repertoire of human kappa light chain sequences
XX      ADO07310-ADO07349 screened for putative catalytic triads. Several genes
XX      encoding such light chains ADO07282-ADO07309 were cloned for use in
XX      recombinant catalytic polypeptides of the invention. These comprise a
XX      human antibody light chain operably joined to a heterologous antibody
XX      heavy chain. The light chain has a serine protease dyad and endopeptidase
XX      activity, and the heavy chain has a predetermined specificity for a
XX      target protein. By joining 2 heterologous human antibody chains, one of
XX      which supplies the catalytic activity to hydrolyse polypeptides and the
XX      other the binding specificity for a target protein, the invention
XX      provides for the construction of a repertoire of proteases with
XX      customised protein substrate specificities of potentially unlimited
XX      number and thus makes possible the effective treatment and/or prevention
XX      of any medical condition attributable to the presence or overexpression
XX      of an identified protein. The invention also provides nucleic acids
XX      encoding the catalytic antibodies (which can be used for gene therapy),
XX      host cells, transgenic non-human animals, and methods of cleaving a
XX      target protein (in vitro or in vivo) using a recombinant catalytic
XX      polypeptide. It also provides a library of recombinant catalytic
XX      polypeptides with altered enzymatic activity, and a method of altering
XX      the enzymatic activity of the recombinant catalytic polypeptides by
XX      mutating at least one complementarity determining region of the heavy
XX      chain. A serine protease triad was not identified in the present
XX      sequence.
XX
XX      Sequence 95 AA:
XX
XX      Query Match      34.0%; Score 88; DB 1; Length 95;
XX      Best Local Similarity 92.6%; Pred. No. 6.6;
XX      Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
XX      387 ELELTGSPATLSVSPGERATLSGRASVSVDLAWYQKQKQAPRLIYAGSTRATGVA 446
XX      1 EIWWTGSPATLSVSPGERATLSGRASVSVDLAWYQKQKQAPRLIYAGSTRATGIP 60
XX
XX      447 RFGSGSGGAFTLTISLQSEDPAYVYCCQYNNWP 481
XX      61 RFGSGSGGTEFTLTISLQSEDPAYVYCCQYNNWP 95
XX
XX      RESULT 20
XX      ID ADO07340 standard; protein; 95 AA.
XX
XX      ADO07340;
XX
XX      15-JUL-2004 (first entry)
XX

```

```

DE      Human antibody L2 light chain variable region.
XX
XX      Catalytic antibody; human; antibody; gene therapy.
XX
XX      Homo sapiens.
XX
XX      Key
XX      Location/Qualifiers
XX      Region
XX      1..23
XX      /label= FR1
XX      /note= "Framework region 1"
XX
XX      Region
XX      24..34
XX      /label= CDR1
XX      /note= "Complementarity determining region 1"
XX
XX      Region
XX      35..49
XX      /label= FR2
XX      /note= "Framework region 2"
XX
XX      Region
XX      50..56
XX      /label= CDR2
XX      /note= "Complementarity determining region 2"
XX
XX      Region
XX      57..88
XX      /label= FR3
XX      /note= "Framework region 3"
XX
XX      Region
XX      89..95
XX      /label= CDR3
XX      /note= "Complementarity determining region 3"
XX
XX      MO2004033658-A2.
XX
XX      22-APR-2004.
XX
XX      09-OCT-2003; 2003WO-US032214.
XX
XX      10-OCT-2002; 2002US-0417979P.
XX
XX      (INTE-) INTEGRIGEN INC.
XX
XX      Smider V, Larrick JW;
XX
XX      WPI; 2004-340921/31.
XX
XX      Novel recombinant catalytic polypeptide useful for cleaving target
PT      proteins or for treating or preventing cancers, comprises a human
PT      antibody light chain operably joined to a heterologous antibody heavy
PT      chain.
XX
XX      Disclosure; Fig 3; 92pp; English.
XX
XX      The present sequence is the light chain variable region of human antibody
XX      116. This is one of repertoire of human kappa light chain sequences
XX      ADO07310-ADO07349 screened for putative catalytic triads. Several genes
XX      encoding such light chains ADO07282-ADO07309 were cloned for use in
XX      recombinant catalytic polypeptides of the invention. These comprise a
XX      human antibody light chain operably joined to a heterologous antibody
XX      heavy chain. The light chain has a serine protease dyad and endopeptidase
XX      activity, and the heavy chain has a predetermined specificity for a
XX      target protein. By joining 2 heterologous human antibody chains, one of
XX      which supplies the catalytic activity to hydrolyse polypeptides and the
XX      other the binding specificity for a target protein, the invention
XX      provides for the construction of a repertoire of proteases with
XX      customised protein substrate specificities of potentially unlimited
XX      number and thus makes possible the effective treatment and/or prevention
XX      of any medical condition attributable to the presence or overexpression
XX      of an identified protein. The invention also provides nucleic acids
XX      encoding the catalytic antibodies (which can be used for gene therapy),
XX      host cells, transgenic non-human animals, and methods of cleaving a
XX      target protein (in vitro or in vivo) using a recombinant catalytic
XX      polypeptide. It also provides a library of recombinant catalytic
XX      polypeptides with altered enzymatic activity, and a method of altering
XX      the enzymatic activity of the recombinant catalytic polypeptides by
XX      mutating at least one complementarity determining region of the heavy
XX      chain. A serine protease triad was not identified in the present
XX      sequence.
XX

```



SQ Sequence 95 AA;

Query Match 34.0%; Score 88; DB 1; Length 95;  
 Best Local Similarity 92.6%; Pred. No. 6.6;  
 Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 387 ELELTGSPATLSVSPGERATLSCRASBSVSSDLAWYQKPGQAPRLIYGASTRATGVPA 446  
 DB 1 EIVWTGSPATLSVSPGERATLSCRASQSVSSNLAWYQKPGQAPRLIYGASTRATGIPA 60  
 OY 447 RFGSGSGAEFTLTISLSQSEDFAVYYCQQTNNMP 481  
 DB 61 RFGSGSGAEFTLTISLSQSEDFAVYYCQQTNNMP 95

RESULT 21

ID ADR43440 standard; peptide; 32 AA.

AC ADR43440;

DT 04-NOV-2004 (first entry)

DE FRL variant #11.

XX antibody; variable light chain; variable heavy chain; Antiallergic;  
 KM Dermatological; Immunosuppressive; IGE; asthma; allergic rhinitis;  
 KM eczema; urticaria; atopic dermatitis; food allergy; CDR; FRL.

OS Unidentified.

PN WO2004070011-A2.

XX 19-AUG-2004.

PF 02-FEB-2004; 2004WO-US002894.

PR 01-FEB-2003; 2003US-0444229P.

PA (TANO-) TANOX INC.

PI Singh S, Foster C, Wu H;

XX WPI; 2004-604433/58.

PT New high affinity human monoclonal antibodies, particularly those  
 directed against isotypic determinants of immunoglobulin E, useful for  
 asthma, allergic rhinitis, eczema, urticaria, atopic dermatitis, or a  
 food allergy.

PS Claim 1; SEQ ID NO 40; 101pp; English.

XX The present invention relates to an antibody comprising a variable light  
 chain region or a variable heavy chain region. The antibody and methods  
 are useful for treating a disorder associated with an abnormally high IGE  
 level, e.g. asthma, allergic rhinitis, eczema, urticaria, atopic  
 dermatitis, or a food allergy. The present sequence represents a FRL  
 CC which demonstrates improved affinity. This sequence is present in the  
 library of variants in the specification.

SQ Sequence 32 AA;

Query Match 11.6%; Score 30; DB 1; Length 32;  
 Best Local Similarity 93.8%; Pred. No. 27;  
 Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 443 GVPARFSGSGAEFTLTISLSQSEDFAVYYC 474  
 DB 1 GVPARFSGSGAEFTLTISLSQSEDFAVYYC 32

RESULT 22

ADR31585

ID ADR31585 standard; peptide; 32 AA.

AC ADR31585;

DT 04-NOV-2004 (first entry)

DE Murine TES-C21 antibody (clone 136) FRK3 variant peptide.

KM Antibody; diagnostic; prophylaxis; therapy; kappa chain variable region;  
 VK; murine; FR; framework region; mutant; mutein; variant.

OS Mus sp.

PN WO2004070010-A2.

XX 19-AUG-2004.

PF 02-FEB-2004; 2004WO-US002892.

PR 01-FEB-2003; 2003US-0444229P.

PA (TANO-) TANOX INC.

PI Singh S, Foster C, Wu H;

XX WPI; 2004-604432/58.

PT Generating a humanized, high affinity antibody from an antibody of  
 PT interest comprises selecting a suitable human template as the framework  
 for the H and L chain variable domains of the high affinity antibody to  
 be made.

PS Example 10; SEQ ID NO 40; 100pp; English.

XX The invention relates to a method for generating a humanised high  
 CC affinity antibody from an antibody of interest. The method involves  
 CC selecting a suitable human template as the framework for the H (heavy)  
 CC and L (light) chain variable (V) domains of the high affinity antibody to  
 CC be made. The method is useful for generating high affinity antibodies  
 CC useful in diagnostics, prophylaxis and treatment of diseases. The present  
 CC sequence is murine anti-human IGE antibody (TES-C21) VK (kappa chain  
 variable region) FR (framework region) variant peptide.

SQ Sequence 32 AA;

Query Match 11.6%; Score 30; DB 1; Length 32;  
 Best Local Similarity 93.8%; Pred. No. 27;  
 Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 443 GVPARFSGSGAEFTLTISLSQSEDFAVYYC 474  
 DB 1 GVPARFSGSGAEFTLTISLSQSEDFAVYYC 32

RESULT 23

ID ADR43442 standard; peptide; 32 AA.

AC ADR43442;

DT 04-NOV-2004 (first entry)

DE FRL variant #13.

XX antibody; variable light chain; variable heavy chain; Antiallergic;  
 KM Dermatological; Immunosuppressive; IGE; asthma; allergic rhinitis;  
 KM eczema; urticaria; atopic dermatitis; food allergy; CDR; FRL.

OS Unidentified.

PN WO2004070011-A2.

PD 19-AUG-2004.  
 XX  
 PF 02-FEB-2004; 2004WO-US002894.  
 XX  
 PR 01-FEB-2003; 2003US-0444229P.  
 XX  
 PA (TANO-) TANOX INC.  
 XX  
 PI Singh S, Foster C, Wu H;  
 XX  
 DR WPI; 2004-604433/58.  
 XX  
 PT New high affinity human monoclonal antibodies, particularly those  
 PT directed against isotypic determinants of immunoglobulin E, useful for  
 PT asthma, allergic rhinitis, eczema, urticaria, atopic dermatitis, or a  
 PT food allergy.  
 XX  
 PS Claim 1; SEQ ID NO 42; 101pp; English.  
 XX  
 CC The present invention relates to an antibody comprising a variable light  
 CC chain region or a variable heavy chain region. The antibody and methods  
 CC are useful for treating a disorder associated with an abnormally high IGE  
 CC level, e.g. asthma, allergic rhinitis, eczema, urticaria, atopic  
 CC dermatitis, or a food allergy. The present sequence represents a FRL  
 CC which demonstrates improved affinity. This sequence is present in the  
 CC library of variants in the specification.  
 CC  
 SQ Sequence 32 AA;

Query Match 11.2%; Score 29; DB 1; Length 32;  
 Best Local Similarity 90.6%; Pred. No. 27;  
 Matches 29; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Qy 443 GVPARFSGSGSGAEFTLTSSLSQSEDPFAYYC 474  
 Db 1 GIPSRFSGSGSGTEFTLTSSLSQSEDPADYYC 32

RESULT 24  
 ADR43441  
 ID ADR43441 standard; peptide; 32 AA.  
 XX  
 AC ADR43441;  
 XX  
 DT 04-NOV-2004 (first entry)  
 XX  
 DE FRL variant #12.  
 XX  
 KM antibody; variable light chain; variable heavy chain; Antiallergic;  
 KM Dermatological; Immunosuppressive; IGE; asthma; allergic rhinitis;  
 KM eczema; urticaria; atopic dermatitis; food allergy; CDR; FRL.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2004070011-A2.  
 XX  
 PD 19-AUG-2004.  
 XX  
 PF 02-FEB-2004; 2004WO-US002894.  
 XX  
 PR 01-FEB-2003; 2003US-0444229P.  
 XX  
 PA (TANO-) TANOX INC.  
 XX  
 PI Singh S, Foster C, Wu H;  
 XX  
 DR WPI; 2004-604433/58.  
 XX  
 PT New high affinity human monoclonal antibodies, particularly those  
 PT directed against isotypic determinants of immunoglobulin E, useful for  
 PT asthma, allergic rhinitis, eczema, urticaria, atopic dermatitis, or a  
 PT food allergy.  
 XX

PS Claim 1; SEQ ID NO 41; 101pp; English.  
 XX  
 CC The present invention relates to an antibody comprising a variable light  
 CC chain region or a variable heavy chain region. The antibody and methods  
 CC are useful for treating a disorder associated with an abnormally high IGE  
 CC level, e.g. asthma, allergic rhinitis, eczema, urticaria, atopic  
 CC dermatitis, or a food allergy. The present sequence represents a FRL  
 CC which demonstrates improved affinity. This sequence is present in the  
 CC library of variants in the specification.  
 CC  
 SQ Sequence 32 AA;

Query Match 11.2%; Score 29; DB 1; Length 32;  
 Best Local Similarity 90.6%; Pred. No. 27;  
 Matches 29; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Qy 443 GVPARFSGSGSGAEFTLTSSLSQSEDPFAYYC 474  
 Db 1 GIPSRFSGSGSGTEFTLTSSLSQSEDPFAYYC 32

RESULT 25  
 ADR31587  
 ID ADR31587 standard; peptide; 32 AA.  
 XX  
 AC ADR31587;  
 XX  
 DT 04-NOV-2004 (first entry)  
 XX  
 DE Murine TES-C21 antibody (clone 13) FRK3 variant peptide.  
 XX  
 KM Antibody; diagnostic; prophylaxis; therapy; kappa chain variable region;  
 KM VK; murine; FR; framework region; mutant; mutein; variant.  
 XX  
 OS Mus sp.  
 OS Synthetic.  
 XX  
 PN WO2004070010-A2.  
 XX  
 PD 19-AUG-2004.  
 XX  
 PF 02-FEB-2004; 2004WO-US002892.  
 XX  
 PR 01-FEB-2003; 2003US-0444229P.  
 XX  
 PA (TANO-) TANOX INC.  
 XX  
 PI Singh S, Foster C, Wu H;  
 XX  
 DR WPI; 2004-604432/58.  
 XX  
 PT Generating a humanized, high affinity antibody from an antibody of  
 PT interest comprises selecting a suitable human template as the framework  
 PT for the H and L chain variable domains of the high affinity antibody to  
 PT be made.  
 XX  
 PS Example 10; SEQ ID NO 42; 100pp; English.  
 XX  
 CC The invention relates to a method for generating a humanised high  
 CC affinity antibody from an antibody of interest. The method involves  
 CC selecting a suitable human template as the framework for the H (heavy)  
 CC and L (light) chain variable (V) domains of the high affinity antibody to  
 CC be made. The method is useful for generating high affinity antibodies  
 CC useful in diagnostics, prophylaxis and treatment of diseases. The present  
 CC sequence is murine anti-human IGE antibody (TES-C21) VK (kappa chain  
 CC variable region) FR (framework region) variant peptide.  
 CC  
 SQ Sequence 32 AA;

Query Match 11.2%; Score 29; DB 1; Length 32;  
 Best Local Similarity 90.6%; Pred. No. 27;  
 Matches 29; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 443 GVPARFSGSGSGAEFTLTISLSLQSEDPAYVYC 474  
 DB 1 GIPARFSGSGSGTEFTLTISLSLQSEDPADYYC 32

RESULT 26  
 ID ADR31586 standard; peptide; 32 AA.

AC ADR31586;

DT 04-NOV-2004 (first entry)

DE Murine TES-C21 antibody (clone 1) FRK3 variant peptide.

KM Antibody; diagnostic; prophylaxis; therapy; kappa chain variable region;  
 VK; murine; FR; framework region; mutant; mutein; variant.

OS Mus sp.  
 OS Synthetic.

XX WO2004070010-A2.

XX 19-AUG-2004.

XX 02-FEB-2004; 2004MO-US002892.

XX 01-FEB-2003; 2003US-0444229P.

XX (TANCO-) TANCOX INC.

XX Singh S, Foster C, Wu H;

XX WPI; 2004-604432/58.

PT Generating a humanized, high affinity antibody from an antibody of  
 PT interest comprises selecting a suitable human template as the framework  
 PT for the H and L chain variable domains of the high affinity antibody to  
 be made.

XX Example 10; SEQ ID NO 41; 100pp; English.

CC The invention relates to a method for generating a humanised high  
 CC affinity antibody from an antibody of interest. The method involves  
 CC selecting a suitable human template as the framework for the H (heavy)  
 CC and L (light) chain variable (V) domains of the high affinity antibodies  
 CC to be made. The method is useful for generating high affinity antibodies  
 CC useful in diagnostics, prophylaxis and treatment of diseases. The present  
 CC sequence is murine anti-human IGF antibody (TES-C21) VK (kappa chain  
 CC variable region) FR (framework region) variant peptide.

XX Sequence 32 AA;

Query Match 11.2%; Score 29; DB 1; Length 32;  
 Best Local Similarity 90.6%; Pred. No. 27;

Matches 29; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 443 GVPARFSGSGSGAEFTLTISLSLQSEDPAYVYC 474

DB 1 GIPARFSGSGSGTEFTLTISLSLQSEDPAYVYC 32

RESULT 27

ID AAU90894 standard; peptide; 26 AA.

AC AAU90894;

DT 18-JUN-2002 (first entry)

DE Insulin/insulin-like growth factor receptor-binding peptide #2850.

XX Cytostatic; antidiabetic; neuroprotective; cerebroprotective;

KM ophthalmological; insulin; receptor; gene therapy; diabetes;  
 KM insulin-like growth factor-1; IGF-1; tumour; prostate; breast;  
 KM diabetic retinopathy; neurological diseases; stroke; diabetic neuropathy.

XX Synthetic.

XX WO200172771-A2.

XX 04-OCT-2001.

XX 29-MAR-2000; 2000MO-US008528.

XX 29-MAR-2000; 2000MO-US008528.

XX (DGI-) DGI BIOTECHNOLOGIES LLC.

XX (NOVO) NOVO NORDISK AS.

XX Beasley J, Blume AJ, Schaeffer L, Pillutla R, Brandt J;  
 PI Brissette R, Spetzler J, Cheng W, Ostergaard S, Mandeckl WS;  
 PI Hansen PH, Ravera M, Hsiao K;

XX WPI; 2002-025774/03.

PT Modulating insulin activity in mammalian cells, for treating e.g.  
 PT diabetes and tumors, comprises using peptides that bind to insulin or  
 PT insulin-like growth factor receptors.

XX Example 5; Fig 33A; 390pp; English.

CC The invention relates to a method of modulating insulin activity in  
 CC mammalian cells by administering a peptide that binds the insulin  
 CC receptor (IR). A composition containing a peptide, optionally expressed  
 CC from gene therapy vectors, that binds to Site 1 of IR and an insulin  
 CC agonist are useful for treating diabetes. Also, peptides that are  
 CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are  
 CC useful for treating insulin-like growth factor (IGF)-sensitive tumours  
 CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1  
 CC receptor agonists are useful for treating neurological diseases,  
 CC including stroke and diabetic neuropathy. The peptides are also useful in  
 CC screening for compounds that bind to IR or IGF-1 receptor, potential  
 CC therapeutics and research reagents. AAU88034-AAU90957 represent IR and/or  
 CC IGF-1 receptor-binding peptides and related amino acid sequences of the  
 CC invention

XX Sequence 26 AA;

Query Match 9.3%; Score 24; DB 1; Length 26;

Best Local Similarity 92.3%; Pred. No. 27;

Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 361 WGCGTLVTVSSGGSGSGSGSGSGSGS 386

DB 1 WGCGTLVTVSSGGSGSGSGSGSGSGS 26

Search completed: September 28, 2005, 10:28:41  
 Job time : 1 secs

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OW protein - protein search, using sw model

Run on: September 28, 2005, 10:30:43 ; Search time 0.001 Seconds  
(without alignments)  
22.274 Million cell updates/sec

Title: US-09-936-702-3  
Perfect score: 259  
Sequence: 1 QVQLLESGAEVKKPGSSSVK.....RYFGQGRLEIKLVPRGSG 259

Scoring table: UNITARY AA  
Gapox 10.0 , Gapext 0.5

Searched: 5 seqs, 86 residues

Total number of hits satisfying chosen parameters: 5

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 5 summaries

Database: ra1.subdb.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	6.9	19	1	PCT-US91-02942-67
2	17	6.6	18	1	US-08-207-169A-7
3	16	6.2	16	1	US-08-331-398A-54
4	16	6.2	16	1	US-08-331-397B-54
5	16	6.2	17	1	US-08-463-163-5

#### ALIGNMENTS

RESULT 1  
PCT-US91-02942-67  
Sequence 67, Application PC/TUS9102942  
GENERAL INFORMATION:  
APPLICANT: ROTHLEIN, ROBERT  
APPLICANT: ADAIR, JOHN R  
APPLICANT: ATHWAL, DILJEET S  
TITLE OF INVENTION: HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY  
NUMBER OF SEQUENCES: 102  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1225 Connecticut Ave. NW Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/02942

FILING DATE: 19910429  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9009549.8  
FILING DATE: 27-APR-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: FOX, SAM L  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 1011.0586600  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 466-0800  
TELEFAX: (202) 833-8716  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-02942-67

Query Match 6.9%; Score 18; DB 1; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.29;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 250 SGAEVKKPGSSSVKSCA 267  
Db 2 SGAEVKKPGSSSVKSCA 19

RESULT 2  
US-08-207-169A-7  
Sequence 7, Application US/08207169A  
Patent No. 5674712  
GENERAL INFORMATION:  
APPLICANT: GRANDI, GUIDO  
APPLICANT: DE FERRA, FRANCESCA  
APPLICANT: TOSI, CLAUDIO  
APPLICANT: TORTORA, ORNELLA  
APPLICANT: CUZZONI, ANNA  
TITLE OF INVENTION: RECOMBINANT VECTOR AND USE THEREOF FOR  
TITLE OF INVENTION: EXOCYLLULAR PREPARATION OF SINGLE MOLECULA ANTIBODIES FROM  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSER: P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/207,169A  
FILING DATE: 08-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Obion, No. 5674712man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 2264-061-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-207-169A-7

Query Match 6.6%; Score 17; DB 1; Length 18;  
Best Local Similarity 94.4%; Pred. No. 1.3;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 369 VTSGGGSGGGSGGGGS 386  
Db 1 VSSGGSGGGSGGGSGGGGS 18

RESULT 3  
US-08-331-398A-54  
Sequence 54, Application US/08311398A  
Patent No. 5608039  
GENERAL INFORMATION:  
APPLICANT: Pastan, Ira  
APPLICANT: Willingham, Mark  
APPLICANT: Fitzgerald, David  
APPLICANT: Brinkmann, Ulrich  
APPLICANT: Pal, Lee  
TITLE OF INVENTION: Single Chain B3 Antibody Fusion Proteins  
TITLE OF INVENTION: and their uses (as amended)  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Street Plaza  
City: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105-1492  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/331,398A  
FILING DATE: 28-OCT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/767,331  
FILING DATE: 30-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/596,289  
FILING DATE: 12-OCT-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunter, Tom  
REGISTRATION NUMBER: 38,498  
REFERENCE/DOCKET NUMBER: 015280-126110US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-331-398A-54

Query Match 6.2%; Score 16; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 SGGGSGGGSGGGSGGGGS 386  
Db 1 SGGGSGGGSGGGSGGGGS 16

RESULT 4  
US-08-331-397B-54  
Sequence 54, Application US/08311397B  
Patent No. 5981726  
GENERAL INFORMATION:  
APPLICANT: Pastan, Ira  
APPLICANT: Benhar, Itai  
TITLE OF INVENTION: Chimeric and Mutationally Stabilized Tumor-Specific Antibody Fragments, Fusion Proteins, and Uses Thereof  
TITLE OF INVENTION: thereof  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Street Plaza  
City: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105-1492  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/331,397B  
FILING DATE: 28-OCT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/767,331  
FILING DATE: 30-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/596,289  
FILING DATE: 12-OCT-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunter, Tom  
REGISTRATION NUMBER: 38,498  
REFERENCE/DOCKET NUMBER: 015280-126120US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-331-397B-54

Query Match 6.2%; Score 16; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 SGGGSGGGSGGGSGGGGS 386  
Db 1 SGGGSGGGSGGGSGGGGS 16

RESULT 5  
US-08-463-163-5  
Sequence 5, Application US/08463163  
Patent No. 5696237  
GENERAL INFORMATION:  
APPLICANT: Fitzgerald, David J.  
APPLICANT: Chaudhary, Vijay K.  
APPLICANT: Pastan, Ira H.  
APPLICANT: Waldmann, Thomas A.  
APPLICANT: Queen, Cary L.  
TITLE OF INVENTION: Recombinant Antibody-Toxin Fusion Protein  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Street Tower

CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94105-1492  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/463,163  
 FILING DATE: 05-JUN-1995  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 06/227,227  
 FILING DATE: 22-JAN-1981  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 06/911,227  
 FILING DATE: 24-SEP-1986  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/341,361  
 FILING DATE: 21-APR-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/865,722  
 FILING DATE: 08-APR-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Weber, Ellen L.  
 REGISTRATION NUMBER: 32,762  
 REFERENCE/DOCKET NUMBER: 015280-12211  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 543-9600  
 TELEFAX: (415) 543-5043  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 17 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-463-163-5

Query Match 6.2%; Score 16; DB 1; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 3.8;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 371 SGGGSGGGSGGGSG 386  
 Db 1 SGGGSGGGSGGGSG 16

Search completed: September 28, 2005, 10:30:44  
 Job time: 1 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 10:33:18 / Search time 0.001 Seconds  
(without alignments)  
304.066 Million cell updates/sec

Title: US-09-936-702-3  
Perfect score: 259  
Sequence: 1 QVQLTSGAEVKKPKSSVAVK.....RYTFGGQTRLEIKLVRGSG 259

Scoring table: UNITARY AA  
Gapop 10.0, Gapept 0.5

Searched: 17 seqs, 1174 residues

Total number of hits satisfying chosen parameters: 17

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

. Listing first 17 summaries

Database: rapb.subdb.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	34.0	95	1 US-10-194-975-84	Sequence 84, Appl
2	88	34.0	95	1 US-10-194-975-85	Sequence 85, Appl
3	88	34.0	95	1 US-10-308-817-31	Sequence 31, Appl
4	88	34.0	95	1 US-10-308-817-32	Sequence 32, Appl
5	88	34.0	95	1 US-10-453-698-31	Sequence 31, Appl
6	88	34.0	95	1 US-10-453-698-32	Sequence 32, Appl
7	88	34.0	95	1 US-10-379-392-95	Sequence 95, Appl
8	88	34.0	95	1 US-10-379-392-96	Sequence 96, Appl
9	88	34.0	95	1 US-10-869-355-22	Sequence 22, Appl
10	88	34.0	95	1 US-10-984-960A-106	Sequence 106, Appl
11	30	11.6	32	1 US-10-923-068-75	Sequence 75, Appl
12	30	11.6	32	1 US-10-923-068-84	Sequence 84, Appl
13	30	11.6	32	1 US-10-984-960A-34	Sequence 34, Appl
14	30	11.6	32	1 US-10-984-960A-70	Sequence 70, Appl
15	30	11.6	32	1 US-10-901-736-40	Sequence 40, Appl
16	29	11.2	32	1 US-10-901-736-41	Sequence 41, Appl
17	29	11.2	32	1 US-10-901-736-42	Sequence 42, Appl

## ALIGNMENTS

RESULT 1  
US-10-194-975-84  
; Sequence 84, Application US/10194975  
; Publication No. US20030039649A1  
; GENERAL INFORMATION:  
; APPLICANT: Foote, Jefferson  
; TITLE OF INVENTION: Super Humanized Antibodies  
; FILE REFERENCE: 501231.01  
; CURRENT APPLICATION NUMBER: US/10/194, 975  
; CURRENT FILING DATE: 2002-10-10  
; PRIOR APPLICATION NUMBER: US 60/305,111

; PRIOR FILING DATE: 2001-07-12  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 84  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-194-975-84

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELRLTQSPATLVSPPGERATLSCRASESYSVDLAWYQKRGQAPRLIIYGASTRATGVP 446  
DB 1 EIVMTQSPATLVSPPGERATLSCRAQSQVSSNLAWYQKRGQAPRLIIYGASTRATGIP 60  
QY 447 RFGSGSGAEFTLTISLSQSEDFAVYYCCQYNNWP 481  
DB 61 RFGSGSGTEFTLTISLSQSEDFAVYYCCQYNNWP 95

RESULT 2  
US-10-194-975-85  
; Sequence 85, Application US/10194975  
; Publication No. US20030039649A1  
; GENERAL INFORMATION:  
; APPLICANT: Foote, Jefferson  
; TITLE OF INVENTION: Super Humanized Antibodies  
; FILE REFERENCE: 501231.01  
; CURRENT APPLICATION NUMBER: US/10/194, 975  
; CURRENT FILING DATE: 2002-10-10  
; PRIOR APPLICATION NUMBER: US 60/305,111  
; PRIOR FILING DATE: 2001-07-12  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 85  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-194-975-85

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELRLTQSPATLVSPPGERATLSCRASESYSVDLAWYQKRGQAPRLIIYGASTRATGVP 446  
DB 1 EIVMTQSPATLVSPPGERATLSCRAQSQVSSNLAWYQKRGQAPRLIIYGASTRATGIP 60  
QY 447 RFGSGSGAEFTLTISLSQSEDFAVYYCCQYNNWP 481  
DB 61 RFGSGSGTEFTLTISLSQSEDFAVYYCCQYNNWP 95

RESULT 3  
US-10-308-817-31  
; Sequence 31, Application US/10308817  
; Publication No. US20030219861A1  
; GENERAL INFORMATION:  
; APPLICANT: Rothen, Russell  
; APPLICANT: Wu, Dayang  
; TITLE OF INVENTION: HYBRID ANTIBODIES  
; FILE REFERENCE: 1087-37  
; CURRENT APPLICATION NUMBER: US/10/308, 817  
; CURRENT FILING DATE: 2002-12-03  
; NUMBER OF SEQ ID NOS: 195  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 31  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: human  
US-10-308-817-31

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELETQSPATLSVSPGERATLSCRASBSVSSDLAWYQKPGQAPRLLIYGASTRATGVA 446  
DB 1 EIVMTQSPATLSVSPGERATLSCRASQSVSSNLAWYQKPGQAPRLLIYGASTRATGIPA 60

QY 447 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 481  
DB 61 RFGSGSGTEFTLTISLQSEDFAVYYCCQYNNWP 95

RESULT 4  
US-10-308-817-32  
; Sequence 32, Application US/10308817  
; Publication No. US20030219861A1  
; GENERAL INFORMATION:  
; APPLICANT: Rother, Russell  
; APPLICANT: Wu, Dayang  
; TITLE OF INVENTION: HYBRID ANTIBODIES  
; FILE REFERENCE: 1087-37  
; CURRENT APPLICATION NUMBER: US/10/308, 817  
; CURRENT FILING DATE: 2002-12-03  
; NUMBER OF SEQ ID NOS: 195  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 32  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: human  
US-10-308-817-32

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELETQSPATLSVSPGERATLSCRASBSVSSDLAWYQKPGQAPRLLIYGASTRATGVA 446  
DB 1 EIVMTQSPATLSVSPGERATLSCRASQSVSSNLAWYQKPGQAPRLLIYGASTRATGIPA 60

QY 447 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 481  
DB 61 RFGSGSGTEFTLTISLQSEDFAVYYCCQYNNWP 95

RESULT 5  
US-10-453-698-31  
; Sequence 31, Application US/10453698  
; Publication No. US20040038308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rother, Russell  
; TITLE OF INVENTION: HYBRID ANTIBODIES  
; FILE REFERENCE: 82 CIP (1087-37 CIP)  
; CURRENT APPLICATION NUMBER: US/10/453, 698  
; CURRENT FILING DATE: 2003-06-03  
; NUMBER OF SEQ ID NOS: 196  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 31  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: human  
US-10-453-698-31

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELETQSPATLSVSPGERATLSCRASBSVSSDLAWYQKPGQAPRLLIYGASTRATGVA 446  
DB 1 EIVMTQSPATLSVSPGERATLSCRASQSVSSNLAWYQKPGQAPRLLIYGASTRATGIPA 60

QY 447 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 481  
DB 61 RFGSGSGTEFTLTISLQSEDFAVYYCCQYNNWP 481

DB 61 RFGSGSGTEFTLTISLQSEDFAVYYCCQYNNWP 95

RESULT 6  
US-10-453-698-32  
; Sequence 32, Application US/10453698  
; Publication No. US20040038308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rother, Russell  
; TITLE OF INVENTION: HYBRID ANTIBODIES  
; FILE REFERENCE: 82 CIP (1087-37 CIP)  
; CURRENT APPLICATION NUMBER: US/10/453, 698  
; CURRENT FILING DATE: 2003-06-03  
; NUMBER OF SEQ ID NOS: 196  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 32  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: human  
US-10-453-698-32

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELETQSPATLSVSPGERATLSCRASBSVSSDLAWYQKPGQAPRLLIYGASTRATGVA 446  
DB 1 EIVMTQSPATLSVSPGERATLSCRASQSVSSNLAWYQKPGQAPRLLIYGASTRATGIPA 60

QY 447 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 481  
DB 61 RFGSGSGTEFTLTISLQSEDFAVYYCCQYNNWP 95

RESULT 7  
US-10-379-392-95  
; Sequence 95, Application US/10379392  
; Publication No. US20040110226A1  
; GENERAL INFORMATION:  
; APPLICANT: Lazar, Gregory Alan  
; APPLICANT: Marshall, John Rudolf  
; APPLICANT: Dahiya, Baabli I.  
; TITLE OF INVENTION: ANTIBODY OPTIMIZATION  
; FILE REFERENCE: A-71386-3 463077-236  
; CURRENT APPLICATION NUMBER: US/10/379, 392  
; CURRENT FILING DATE: 2003-03-03  
; PRIOR APPLICATION NUMBER: US 60/360, 843  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 60/384, 197  
; PRIOR FILING DATE: 2002-05-29  
; NUMBER OF SEQ ID NOS: 184  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 95  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-379-392-95

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELETQSPATLSVSPGERATLSCRASBSVSSDLAWYQKPGQAPRLLIYGASTRATGVA 446  
DB 1 EIVMTQSPATLSVSPGERATLSCRASQSVSSNLAWYQKPGQAPRLLIYGASTRATGIPA 60

QY 447 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 481  
DB 61 RFGSGSGTEFTLTISLQSEDFAVYYCCQYNNWP 95

RESULT 8  
US-10-379-392-96  
; Sequence 96, Application US/10379392  
; Publication No. US20040110226A1  
; GENERAL INFORMATION:  
; APPLICANT: Lazar, Gregory Alan  
; APPLICANT: Desjardais, John Rudolf  
; APPLICANT: Marshall, Shannon Alicia  
; APPLICANT: Dahiyat, Basail I.  
; TITLE OF INVENTION: ANTIBODY OPTIMIZATION  
; FILE REFERENCE: A-71386-3 463077-236  
; CURRENT APPLICATION NUMBER: US/10/379,392  
; PRIOR FILING DATE: 2003-03-03  
; PRIOR APPLICATION NUMBER: US 60/360,843  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 60/384,197  
; PRIOR FILING DATE: 2002-05-29  
; NUMBER OF SEQ ID NOS: 184  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 96  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-379-392-96

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Db 1 EIWMTGSPATLSVSPGERATLSCRASVSVDLAWYQKQKQAPRLIIYGASTRATGVA 446  
387 EELTQSPATLSVSPGERATLSCRASVSVDLAWYQKQKQAPRLIIYGASTRATGVA 446  
1 EIWMTGSPATLSVSPGERATLSCRASVSVDLAWYQKQKQAPRLIIYGASTRATGIPA 60  
Qy 447 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNMP 481  
61 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNMP 95  
Db 61 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNMP 95

RESULT 9  
US-10-869-355-22  
; Sequence 22, Application US/10869355  
; Publication No. US20050048578A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Dongxiao  
; TITLE OF INVENTION: METHODS OF SCREENING FOR MONOCLONAL  
; FILE REFERENCE: EPT-007  
; CURRENT APPLICATION NUMBER: US/10/869,355  
; CURRENT FILING DATE: 2004-06-15  
; PRIOR APPLICATION NUMBER: 60/483,391  
; PRIOR FILING DATE: 2003-06-26  
; PRIOR APPLICATION NUMBER: 60/484,185  
; PRIOR FILING DATE: 2003-06-30  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Oryctolagus cuniculus  
US-10-869-355-22

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Db 1 EIWMTGSPATLSVSPGERATLSCRASVSVDLAWYQKQKQAPRLIIYGASTRATGVA 446  
387 EELTQSPATLSVSPGERATLSCRASVSVDLAWYQKQKQAPRLIIYGASTRATGVA 446  
1 EIWMTGSPATLSVSPGERATLSCRASVSVDLAWYQKQKQAPRLIIYGASTRATGIPA 60  
Qy 447 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNMP 481  
61 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNMP 95  
Db 61 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNMP 95

RESULT 10  
US-10-984-960A-106  
; Sequence 106, Application US/10984960A  
; Publication No. US20050142137A1  
; GENERAL INFORMATION:  
; APPLICANT: Gallo, Michael  
; APPLICANT: Chui, Daniel  
; APPLICANT: Zhong, Haihong  
; APPLICANT: Aza, Gulshan  
; APPLICANT: LaRoche, William J.  
; TITLE OF INVENTION: ANTIBODIES AGAINST SECRETORYLEUKOCYTE PROTEASE INHIBITOR  
; FILE REFERENCE: Cura 970  
; CURRENT APPLICATION NUMBER: US/10/984,960A  
; CURRENT FILING DATE: 2004-11-08  
; PRIOR APPLICATION NUMBER: 60/518,275  
; PRIOR FILING DATE: 2003-11-07  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: CuraSeqlet version 0.1  
; SEQ ID NO 106  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-984-960A-106

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Db 1 EIWMTGSPATLSVSPGERATLSCRASVSVDLAWYQKQKQAPRLIIYGASTRATGVA 446  
387 EELTQSPATLSVSPGERATLSCRASVSVDLAWYQKQKQAPRLIIYGASTRATGVA 446  
1 EIWMTGSPATLSVSPGERATLSCRASVSVDLAWYQKQKQAPRLIIYGASTRATGIPA 60  
Qy 447 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNMP 481  
61 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNMP 95  
Db 61 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNMP 95

RESULT 11  
US-10-923-068-75  
; Sequence 75, Application US/10923068  
; Publication No. US20050042664A1  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Herren  
; APPLICANT: Damschroder, Melissa  
; TITLE OF INVENTION: HUMANIZATION OF ANTIBODIES  
; FILE REFERENCE: AB600US  
; CURRENT APPLICATION NUMBER: US/10/923,068  
; CURRENT FILING DATE: 2004-08-20  
; NUMBER OF SEQ ID NOS: 518  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 75  
; LENGTH: 32  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-923-068-75

Query Match 11.6%; Score 30; DB 1; Length 32;  
Best Local Similarity 93.8%; Pred. No. 16;  
Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Db 1 GIPARFSGSGAEFTLTISLQSEDFAVYYC 32  
443 GIPARFSGSGAEFTLTISLQSEDFAVYYC 474  
Qy 443 GIPARFSGSGAEFTLTISLQSEDFAVYYC 474  
1 GIPARFSGSGAEFTLTISLQSEDFAVYYC 32  
Db 1 GIPARFSGSGAEFTLTISLQSEDFAVYYC 32

RESULT 12  
US-10-923-068-84  
; Sequence 84, Application US/10923068  
; Publication No. US20050042664A1  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Herren

```

; APPLICANT: Dall'Acqua, William
; APPLICANT: Damechroder, Melissa
; TITLE OF INVENTION: HUMANIZATION OF ANTIBODIES
; FILE REFERENCE: AB600US
; CURRENT APPLICATION NUMBER: US/10/923,068
; CURRENT FILING DATE: 2004-08-20
; NUMBER OF SEQ ID NOS: 518
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-923-068-84

Query Match      11.6%; Score 30; DB 1; Length 32;
Best Local Similarity 93.8%; Pred. No. 16;
Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      443 GVPARFSGSGGAEFTLTITSSLOSEDPFAYYC 474
DB      1 GIPARFSGSGGTEFTLTITSSLOSEDPFAYYC 32

RESULT 13
US-10-984-960A-34
; Sequence 34, Application US/10984960A
; Publication No. US20050142137A1
; GENERAL INFORMATION:
; APPLICANT: Gallo, Michael
; APPLICANT: Chui, Daniel
; APPLICANT: Zhong, Haihong
; APPLICANT: Ara, Gulshan
; APPLICANT: LaRochelelle, William J.
; TITLE OF INVENTION: ANTIBODIES AGAINST SECRETORYLEUKOCYTE PROTEASE INHIBITOR
; FILE REFERENCE: Cura 970
; CURRENT APPLICATION NUMBER: US/10/984,960A
; CURRENT FILING DATE: 2004-11-08
; PRIOR APPLICATION NUMBER: 60/518,275
; PRIOR FILING DATE: 2003-11-07
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 34
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-984-960A-34

Query Match      11.6%; Score 30; DB 1; Length 32;
Best Local Similarity 93.8%; Pred. No. 16;
Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      443 GVPARFSGSGGAEFTLTITSSLOSEDPFAYYC 474
DB      1 GIPARFSGSGGTEFTLTITSSLOSEDPFAYYC 32

RESULT 14
US-10-984-960A-70
; Sequence 70, Application US/10984960A
; Publication No. US20050142137A1
; GENERAL INFORMATION:
; APPLICANT: Gallo, Michael
; APPLICANT: Chui, Daniel
; APPLICANT: Zhong, Haihong
; APPLICANT: Ara, Gulshan
; APPLICANT: LaRochelelle, William J.
; TITLE OF INVENTION: ANTIBODIES AGAINST SECRETORYLEUKOCYTE PROTEASE INHIBITOR
; FILE REFERENCE: Cura 970
; CURRENT APPLICATION NUMBER: US/10/984,960A
; CURRENT FILING DATE: 2004-11-08
; PRIOR APPLICATION NUMBER: 60/518,275
; PRIOR FILING DATE: 2003-11-07
; NUMBER OF SEQ ID NOS: 112
```

```

; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 70
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-984-960A-70

Query Match      11.6%; Score 30; DB 1; Length 32;
Best Local Similarity 93.8%; Pred. No. 16;
Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      443 GVPARFSGSGGAEFTLTITSSLOSEDPFAYYC 474
DB      1 GIPARFSGSGGTEFTLTITSSLOSEDPFAYYC 32

RESULT 15
US-10-901-736-40
; Sequence 40, Application US/10901736
; Publication No. US20050169909A1
; GENERAL INFORMATION:
; APPLICANT: TANOX, INC.
; APPLICANT: SINGH, Sanjaya
; APPLICANT: HUANG, Danyang
; APPLICANT: FUNG, Sek Chung
; TITLE OF INVENTION: Identification of Unique, High Affinity Igs Epitopes
; FILE REFERENCE: TNX-1030
; CURRENT APPLICATION NUMBER: US/10/901,736
; CURRENT FILING DATE: 2004-07-29
; PRIOR APPLICATION NUMBER: 60/444,229
; PRIOR FILING DATE: 2003-02-01
; PRIOR APPLICATION NUMBER: PCT/US04/02892
; PRIOR FILING DATE: 2004-02-02
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 32
; TYPE: PRT
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: FRU3 VARIANT 136
US-10-901-736-40

Query Match      11.6%; Score 30; DB 1; Length 32;
Best Local Similarity 93.8%; Pred. No. 16;
Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      443 GVPARFSGSGGAEFTLTITSSLOSEDPFAYYC 474
DB      1 GIPARFSGSGGTEFTLTITSSLOSEDPFAYYC 32

RESULT 16
US-10-901-736-41
; Sequence 41, Application US/10901736
; Publication No. US20050169909A1
; GENERAL INFORMATION:
; APPLICANT: TANOX, INC.
; APPLICANT: SINGH, Sanjaya
; APPLICANT: HUANG, Danyang
; APPLICANT: FUNG, Sek Chung
; TITLE OF INVENTION: Identification of Unique, High Affinity Igs Epitopes
; FILE REFERENCE: TNX-1030
; CURRENT APPLICATION NUMBER: US/10/901,736
; CURRENT FILING DATE: 2004-07-29
; PRIOR APPLICATION NUMBER: 60/444,229
; PRIOR FILING DATE: 2003-02-01
; PRIOR APPLICATION NUMBER: PCT/US04/02892
; PRIOR FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: PCT/US04/02894
; PRIOR FILING DATE: 2004-02-02
```

; NUMBER OF SEQ ID NOS: 77  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 41  
 ; LENGTH: 32  
 ; TYPE: PRT  
 ; ORGANISM: ARTIFICIAL  
 ; FEATURE:  
 ; OTHER INFORMATION: FRL3 VARIANT 1  
 US-10-901-736-41

Query Match 11.2%; Score 29; DB 1; Length 32;  
 Best Local Similarity 90.6%; Pred. No. 16;  
 Matches 29; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 443 GVPARFSGSGCAEFTLTISLSQSEDPFAYYYC 474  
 DB 1 GIPARFSGSGGTEFTLTISLSQSEDPFAYYYC 32

RESULT 17  
 US-10-901-736-42  
 ; Sequence 42, Application US/10901736  
 ; Publication No. US20050169903A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: TANOX, INC.  
 ; APPLICANT: SINGH, Sanjaya  
 ; APPLICANT: HUANG, Danyang  
 ; APPLICANT: FUNG, Sek Chung  
 ; TITLE OF INVENTION: Identification of Unique, High Affinity Ige Epitopes  
 ; FILE REFERENCE: TXN-1030  
 ; CURRENT APPLICATION NUMBER: US/10/901,736  
 ; CURRENT FILING DATE: 2004-07-29  
 ; PRIOR APPLICATION NUMBER: 60/444,229  
 ; PRIOR FILING DATE: 2003-02-01  
 ; PRIOR APPLICATION NUMBER: PCT/US04/02892  
 ; PRIOR FILING DATE: 2004-02-02  
 ; PRIOR APPLICATION NUMBER: PCT/US04/02894  
 ; PRIOR FILING DATE: 2004-02-02  
 ; NUMBER OF SEQ ID NOS: 77  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 42  
 ; LENGTH: 32  
 ; TYPE: PRT  
 ; ORGANISM: ARTIFICIAL  
 ; FEATURE:  
 ; OTHER INFORMATION: FRL3 VARIANT 13  
 US-10-901-736-42

Query Match 11.2%; Score 29; DB 1; Length 32;  
 Best Local Similarity 90.6%; Pred. No. 16;  
 Matches 29; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 443 GVPARFSGSGCAEFTLTISLSQSEDPFAYYYC 474  
 DB 1 GIPARFSGSGGTEFTLTISLSQSEDPFAYYYC 32

Search completed: September 28, 2005, 10:33:18  
 Job time : 0.001 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: September 28, 2005, 10:42:58 ; Search time 0.001 Seconds  
(without alignments)  
21.756 Million cell updates/sec

Title: US-09-936-702-3  
Perfect score: 259  
Sequence: 1 QVQLSGAEVKKKSSVKV.....RYFGQSTRLEIKLVPRSG 259

Scoring table: UNITARY AA  
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 84 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : tpr.subdb.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	30.9	84	1 S34099	

## ALIGNMENTS

RESULT 1  
S34099

Ig kappa chain V region - human  
C:Species: Homo sapiens (man)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C:Accession: S34099  
R:Wagner, S.D.; Luzatco, L.  
Eur. J. Immunol. 23, 391-397, 1993  
A:Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distributed A:Reference number: S34076, MUID:93170387, PMID:8436174  
A:Accession: S34099  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-84 <WAG>  
A:Cross-references: UNIPROT:Q9UL83; EMBL:X67183  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:5-79/Domain: immunoglobulin homology <IMM>

Query Match 30.9%; Score 80; DB 1; Length 84;  
Best Local Similarity 95.2%; Pred. No. 0;  
Matches 80; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 398 SVSPGERATISCRASESVSDLAHYQQKPGQAPRLITYGASTRATGVPAFSGSGGAEEF 457  
Db 1 SVSPGERATISCRASESVSDLAHYQQKPGQAPRLITYGASTRATGVPAFSGSGGAEEF 60

QY 458 TLTITSSIQSEDFAVYYCCQYNMNP 481  
Db 61 TLTITSSIQSEDFAVYYCCQYNMNP 84

Search completed: September 28, 2005, 10:42:59  
Job time : 0.001 secs

**THIS PAGE BLANK (USPTO)**



Qy	1	ATGAAACCGGGAGTCCCTTTTAAAGGACATTGCTTCTGTGCTGTGACAATTGGCGCTCTCCCA	60
Db	79	ATGAACCGGGAGTCCCTTTTAAAGGACATTGCTTCTGTGCTGTGACAATTGGCGCTCTCCCA	138
Qy	61	GCAGCCACTCAGAGGAAAAGAAAGTGTGTGTGGGCAAAAAAGGGAGTACAGTGGAACTGACC	120
Db	139	GCAGCCACTCAGAGGAAAAGAAAGTGTGTGTGGGCAAAAAAGGGAGTACAGTGGAACTGACC	198
Qy	121	TGTACAGCTTCCCAAGAGAGCATCAATTTCCACTGGAAAAATCTCAACACAGATTAAG	180
Db	199	TGTACAGCTTCCCAAGAGAGCATCAATTTCCACTGGAAAAATCTCAACACAGATTAAG	258
Qy	181	ATTCTGGGAAATCAGGGGCTCTTCTTAACTTAAAGTGCATCCAAAGCTGAATGATCGGCT	240
Db	259	ATTCTGGGAAATCAGGGGCTCTTCTTAACTTAAAGTGCATCCAAAGCTGAATGATCGGCT	318
Qy	241	GACTCAAGAAAGGCTTTGGGACCAAGGAAATTTCCCTCGATCATCAAGATCTTTAAG	300
Db	319	GACTCAAGAAAGGCTTTGGGACCAAGGAAATTTCCCTCGATCATCAAGATCTTTAAG	378
Qy	301	ATAGAAGATCTGATATCTTACTCTGTGTAAGTGAAGACAGAAAGAGAGAGTGCATTG	360
Db	379	ATAGAAGATCTGATATCTTACTCTGTGTAAGTGAAGACAGAAAGAGAGAGTGCATTG	438
Qy	361	CTTAGTGTTCGATTGACTGCGCAACTTGACACCACTGCTTTCAGGGGCAAGGCTGACC	420
Db	439	CTTAGTGTTCGATTGACTGCGCAACTTGACACCACTGCTTTCAGGGGCAAGGCTGACC	498
Qy	421	CTGACCTTGGAGAGCCCCCTGTGTGTAGTACCCCTCAGTGCACAATAGAGATCCAAAGGGT	480
Db	499	CTGACCTTGGAGAGCCCCCTGTGTGTAGTACCCCTCAGTGCACAATAGAGATCCAAAGGGT	558

Qy	481	AAAAACATACAGGGGGGGGAGAACCCCTCTCCGGTCTCAGCTGGAGCTCAGAGTAAATGAGG	54.0
Db	559	AAAAACATACAGGGGGGGGAGAACCCCTCTCCGGTCTCAGCTGGAGCTCAGAGTAAATGAGG	61.8
Qy	541	ACCTGGACATGACATGCTCTTTGCAGAACCGAGAAAGCTGAGCTCAAAATTAAGCATCGT	60.0
Db	619	ACCTGGACATGACATGCTCTTTGCAGAACCGAGAAAGCTGAGCTCAAAATTAAGCATCGT	67.8
Qy	601	GTGCTAGCTTTCAGAGGCTCTCCGAGGCTGCGCGTAGTGCGGGAGGCGG	64.9
Db	679	GTGCTAGCTTTCAGAGGCTCTCCGAGCATAGTTATTAAGAAAGAGGCG	72.7

## RESULT 2

```

PCT-US92-08090-4
Sequence 4, Application PC/TUS9208090
GENERAL INFORMATION:
APPLICANT: Tang, J. N.
TITLE OF INVENTION: Fusion Protein Genes for Treatment of
TITLE OF INVENTION: Aids
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/08090
FILING DATE: 19920922
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patricia L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMAF129
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-572-6508
TELEFAX: 404-572-6555
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1415 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: epithelial
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..6
OTHER INFORMATION: /note= "Restriction site"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1275..1280
OTHER INFORMATION: /note= "Restriction site"
NAME/KEY: misc_feature
LOCATION: 1410..1415
OTHER INFORMATION: /note= "Restriction site"
PCT-US92-08090-4
Query Match 43.4%; Score 624.4; DB 5;
Best Local Similarity 99.8%; Pred. No. 5; 5e-138; Length 1415;

```

	Matches	625:	Conservative	0:	Mismatches	1:	Indels	0:	Gaps	0:
QY		1	ATGAACGGGGAGTCCCTTTTAGGCACTTGCTTCTGTGCTGCGCACTGGCGCTCTCCCA	60						
Db		81	ATGAACCGGGGAGTCCCTTTTAGGCACTTGCTTCTGTGCTGCGCACTGGCGCTCTCCCA	140						
QY		61	GCAGCACTCAGGGGAAAGAAAGTGGTGTGGGCAAAAAGGGGATACAGTGGAACTGAC	120						
Db		141	GCAGCACTCAGGGGAAAGAAAGTGGTGTGGGCAAAAAGGGGATACAGTGGAACTGAC	200						
QY		121	TGTACAGCTTCCAGAGAGAGACATACATTCCACTGGAAAACTCCAACCATTAAG	180						
Db		201	TGTACAGCTTCCAGAGAGAGACATACATTCCACTGGAAAACTCCAACCATTAAG	260						
QY		181	ATTCTGGGAATCACTGGGCTCCTTTCTTAATAAGGTCCATCCAACTGAATGATGGCT	240						
Db		261	ATTCTGGGAATCACTGGGCTCCTTTCTTAATAAGGTCCATCCAACTGAATGATGGCT	320						
QY		241	GACTCAAGAAAGCCTTTGGGACCAAGAAATCTCCCTGATCATCAGATCTTTAAG	300						
Db		321	GACTCAAGAAAGCCTTTGGGACCAAGAAATCTCCCTGATCATCAGATCTTTAAG	380						
QY		301	ATGAGAGCTCAGATACTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCATTG	360						
Db		381	ATGAGAGCTCAGATACTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCATTG	440						
QY		361	CTAGGTTCCGATTGACCTGCCACTCTGACACCAACCTGCTTCAAGGGCAGAGCTGAC	420						
Db		441	CTAGGTTCCGATTGACCTGCCACTCTGACACCAACCTGCTTCAAGGGCAGAGCTGAC	500						
QY		421	CTGACCTTGGAGAGCCCCCTGTAGTAGGACCCCTCAGTGCATGTAGGAGTCCAAGGGT	480						
Db		501	CTGACCTTGGAGAGCCCCCTGTAGTAGGACCCCTCAGTGCATGTAGGAGTCCAAGGGT	560						
QY		481	AAAAACATACAGGGGGGGGAAACCTCTCCGTCTCAGCTGGAGCTCCAAGATTGTGGC	540						
Db		561	AAAAACATACAGGGGGGGGAAACCTCTCCGTCTCAGCTGGAGCTCCAAGATTGTGGC	620						
QY		541	ACCTGSAATCACACTGTCTTGCAGAACCAAGAAAGGTGAGATTCAAATATAGACATCGTG	600						
Db		621	ACCTGSAATCACACTGTCTTGCAGAACCAAGAAAGGTGAGATTCAAATATAGACATCGTG	680						
QY		601	GTGCTAGCTTTCAGAGAGGCTCCGG	626						
Db		681	GTGCTAGCTTTCAGAGAGGCTCCAG	706						

### RESULT 3

```

PCT-US92-08090-3
; Sequence 3, Application PC/US9208090
; GENERAL INFORMATION:
; APPLICANT: Tang, J. N.
; TITLE OF INVENTION: Fusion Protein Genes for Treatment of
; TITLE OF INVENTION: Aids
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08090
; FILING DATE: 19920922
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patricia L.

```

```

;
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRP129
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1421 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Epithelial
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1275..1280
; OTHER INFORMATION: /note= "Restriction site"
;
; NAME/KEY: misc_feature
; LOCATION: 1..6
; OTHER INFORMATION: /note= "Restriction site"
;
; NAME/KEY: misc_feature
; LOCATION: 1275..1280
; OTHER INFORMATION: /note= "Restriction site"
;
; NAME/KEY: misc_feature
; LOCATION: 1416..1421
; OTHER INFORMATION: /note= "Restriction site"
;
PCT-US92-08090-3

```

Query Match 43.4%; Score 624.4; DB 5; Length 1421;

Best Local Similarity 99.8%; Pred. No. 5.5e-138;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGCACTGGGCGCTCCCTCCA 60
DB 81 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGCACTGGGCGCTCCCTCCA 140
QY 61 GCAGCCACTCAGGGAAGAGTGTGCTGGGCAAAAAGGGGATACAGTGAAGTACC 120
DB 141 GCAGCCACTCAGGGAAGAGTGTGCTGGGCAAAAAGGGGATACAGTGAAGTACC 200
QY 121 TGTACAGCTTCCCAAGAAAGAGCATATTCATCTGGAAGAAAGTCCCAAGATTAAG 180
DB 201 TGTACAGCTTCCCAAGAAAGAGCATATTCATCTGGAAGAAAGTCCCAAGATTAAG 260
QY 181 ATTCTGGGAATCAGGGGCTCTTCTTAACCTAAAGTCCATCCAGCTGAATGCGGCT 240
DB 261 ATTCTGGGAATCAGGGGCTCTTCTTAAAGTCCATCCAGCTGAATGCGGCT 320
QY 241 GACTCAAGAAAGCTTTTGGGCAAGAAATTCCTCCCTGATCATCAAGATCTTAAAG 300
DB 321 GACTCAAGAAAGCTTTTGGGCAAGAAATTCCTCCCTGATCATCAAGATCTTAAAG 380
QY 301 ATAGAAGATCAGATATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 360
DB 381 ATAGAAGATCAGATATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 440
QY 361 CTAGTGTGGATTGATGCTGCACTGCAACCTGCTCTTCAAGGGGAGAGCTGAGC 420
DB 441 CTAGTGTGGATTGATGCTGCACTGCAACCTGCTCTTCAAGGGGAGAGCTGAGC 500
QY 421 CTGACCTTGAAGAGCCCTGCTGATAGACCCCTCAGTCAATGTAGAGTCCAGGGGT 480
DB 501 CTGACCTTGAAGAGCCCTGCTGATAGACCCCTCAGTCAATGTAGAGTCCAGGGGT 560
QY 481 AAAAATATCAGGGGGGAAAGACCTCTCCGTCTCAGTGAAGTCCAGATTAAGTGC 540
DB 561 AAAAATATCAGGGGGGAAAGACCTCTCCGTCTCAGTGAAGTCCAGATTAAGTGC 620
QY 541 AACTGAGATGACCTGTCTTGGAGAACCAAGAGAGTGAAGTCAAAATAGACATCGG 600

```

```

DB 621 AACTGAGATGACCTGTCTTGGAGAACCAAGAGAGTGAAGTCAAAATAGACATCGG 680
QY 601 GTGCTAGCTTTCAGAAAGGCTCCCG 626
DB 681 GTGCTAGCTTTCAGAAAGGCTCCCG 706

```

RESULT 4  
PCT-US92-08090-2

```

; Sequence 2, Application PC/TUS9208090
; GENERAL INFORMATION:
; APPLICANT: Tang, J. N.
; TITLE OF INVENTION: Fusion Protein Genes for Treatment of
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08090
; FILING DATE: 19920922
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRP129
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1448 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Epithelial
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..6
; OTHER INFORMATION: /note= "Restriction site"
;
; NAME/KEY: misc_feature
; LOCATION: 1275..1280
; OTHER INFORMATION: /note= "Restriction site"
;
; NAME/KEY: misc_feature
; LOCATION: 1444..1448
; OTHER INFORMATION: /note= "Restriction site"
;
PCT-US92-08090-2

```

Query Match 43.4%; Score 624.4; DB 5; Length 1448;

Best Local Similarity 99.8%; Pred. No. 5.5e-138;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGCACTGGGCGCTCCCTCCA 60
DB 81 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGCACTGGGCGCTCCCTCCA 140
QY 61 GCAGCCACTCAGGGAAGAGTGTGCTGGGCAAAAAGGGGATACAGTGAAGTACC 120

```

```
Db 141 GCAGCCACTCAGGAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAACCTGACC 200
Qy 121 TGTACAGCTTCCCGAAGAAAGACATCAATTCCACTGGAATAAACTCCAACTCAATTAAG 180
Db 201 TGTACAGCTTCCCGAAGAAAGACATCAATTCCACTGGAATAAACTCCAACTCAATTAAG 260
Qy 181 ATTCTGGGAAATCAGGGCTCCTTCTTAACCTAAAGTCCATCCAACTGAATGATCGGCT 240
Db 261 ATTCTGGGAAATCAGGGCTCCTTCTTAACCTAAAGTCCATCCAACTGAATGATCGGCT 320
Qy 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 300
Db 321 GACTCAAGAAAGACCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 380
Qy 301 ATGAAAGCTCAGATCTTACATCTGGAAGTGAAGGACGAGAGAGGAGGATGCAATTG 360
Db 381 ATGAAAGCTCAGATCTTACATCTGGAAGTGAAGGACGAGAGAGGAGGATGCAATTG 440
Qy 361 CTAGTGTTCGGATTGACTGCGCACTGACACCCACCTGCTTCAAGGGCAGAGCTGACC 420
Db 441 CTAGTGTTCGGATTGACTGCGCACTGACACCCACCTGCTTCAAGGGCAGAGCTGACC 500
Qy 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGAATGAGAGTCCAAAGGGGT 480
Db 501 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGAATGAGAGTCCAAAGGGGT 560
Qy 481 AAAAACAATACAGGGGGGGAAGACCCCTGCTGCTCAGCTGAGGCTCAGAGATAGTGGC 540
Db 561 AAAAACAATACAGGGGGGGAAGACCCCTGCTGCTCAGCTGAGGCTCAGAGATAGTGGC 620
Qy 541 ACCTGACATGCACTGCTTTCAGAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 600
Db 621 ACCTGACATGCACTGCTTTCAGAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 680
Qy 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626
Db 681 GTGCTAGCTTTCAGAAAGGCTCCAG 706
```

```
RESULT 5
US-09-517-605-7
; Sequence 7, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geljtenbeck, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1742
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-517-605-7
```

```
Query Match 43.4%; Score 624.4; DB 3; Length 1742;
Best Local Similarity 99.8%; Pred. No. 5.8e-138;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGAACCGGGAGTCCCTTTAGCACTTGTGTGCTGCAACTGGCGCTCTCCCA 60
Db 76 ATGAACCGGGAGTCCCTTTAGCACTTGTGTGCTGCAACTGGCGCTCTCCCA 135
Qy 61 GCAGCCACTCAGGAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAACCTGACC 120
Db 136 GCAGCCACTCAGGAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAACCTGACC 195
```

```
Qy 121 TGTACAGCTTCCCGAAGAAAGACATCAATTCCACTGGAATAAACTCCAACTCAATTAAG 180
Db 196 TGTACAGCTTCCCGAAGAAAGACATCAATTCCACTGGAATAAACTCCAACTCAATTAAG 255
Qy 181 ATTCTGGGAAATCAGGGCTCCTTCTTAACCTAAAGTCCATCCAACTGAATGATCGGCT 240
Db 256 ATTCTGGGAAATCAGGGCTCCTTCTTAACCTAAAGTCCATCCAACTGAATGATCGGCT 315
Qy 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 300
Db 316 GACTCAAGAAAGACCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 375
Qy 301 ATGAAAGCTCAGATCTTACATCTGGAAGTGAAGGACGAGAGAGGAGTGCATTTG 360
Db 376 ATGAAAGCTCAGATCTTACATCTGGAAGTGAAGGACGAGAGAGGAGTGCATTTG 435
Qy 361 CTAGTGTTCGGATTGACTGCGCACTGACACCCACCTGCTTCAAGGGCAGAGCTGACC 420
Db 436 CTAGTGTTCGGATTGACTGCGCACTGACACCCACCTGCTTCAAGGGCAGAGCTGACC 495
Qy 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGAATGAGAGTCCAAAGGGGT 480
Db 496 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGAATGAGAGTCCAAAGGGGT 555
Qy 481 AAAAACAATACAGGGGGGGAAGACCCCTGCTGCTCAGCTGAGGCTCAGAGATAGTGGC 540
Db 556 AAAAACAATACAGGGGGGGAAGACCCCTGCTGCTCAGCTGAGGCTCAGAGATAGTGGC 615
Qy 541 ACCTGACATGCACTGCTTTCAGAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 600
Db 616 ACCTGACATGCACTGCTTTCAGAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 675
Qy 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626
Db 676 GTGCTAGCTTTCAGAAAGGCTCCAG 701
```

```
RESULT 6
US-09-023-655-1013
; Sequence 1013, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Sellhammer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: word Perfect 6.1 for Windows/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
```

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 855-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 1013:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1742 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: GENBANK  
 CLONE: 9179143  
 US-09-023-655-1013

Query Match 43.4%; Score 624.4; DB 4; Length 1742;  
 Best Local Similarity 99.8%; Pred. No. 5.8e-138;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGCACTGCGCTCCCTCCCA 60  
 76 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGCACTGCGCTCCCTCCCA 135  
 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCCAAAAGGGGATACAGTGAAGTACC 120  
 136 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCCAAAAGGGGATACAGTGAAGTACC 195  
 121 TGTACAGCTTCCAGAAAGAGCATACATTCCTGGAATAAATCTCCACCAATTAAG 180  
 196 TGTACAGCTTCCAGAAAGAGCATACATTCCTGGAATAAATCTCCACCAATTAAG 255  
 181 ATTCTGGGAATCAGGGCTCCTTTTAAGTCAATCAAGCTGCAAGTGAATGCGCGT 240  
 256 ATTCTGGGAATCAGGGCTCCTTTTAAGTCAATCAAGCTGCAAGTGAATGCGCGT 315  
 241 GACTCAAGAGAGGCTTTGGAGCAAGAAATTCCTCCCTGATCATCAAGATCTTAAG 300  
 316 GACTCAAGAGAGGCTTTGGAGCAAGAAATTCCTCCCTGATCATCAAGATCTTAAG 375  
 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGAGACCAAGAGAGAGGTGCAATTG 360  
 376 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGAGACCAAGAGAGAGGTGCAATTG 435  
 361 CTAGGTTCGATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 420  
 436 CTAGGTTCGATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 495  
 421 CTAGGTTCGATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 480  
 496 CTAGGTTCGATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 555  
 481 AAAAATATACAGGGGGGAAAGACCTCTCTGCTGCTCAAGTGAAGTCAAGTGAAGT 540  
 556 AAAAATATACAGGGGGGAAAGACCTCTCTGCTGCTCAAGTGAAGTCAAGTGAAGT 615  
 541 ACCTGAGATGACATCTGCTGGAAGAACCAAGAGAGTGAATTAATAATCAATCGG 600  
 616 ACCTGAGATGACATCTGCTGGAAGAACCAAGAGAGTGAATTAATAATCAATCGG 675  
 601 GTGCTAGCTTTCAGAAAGGCTCCCG 626  
 676 GTGCTAGCTTTCAGAAAGGCTCCCG 701

RESULT 7  
 PCT-US92-08090-1  
 Sequence 1, Application PC/TUS9208090

GENERAL INFORMATION:  
 APPLICANT: Tang, J. N.  
 TITLE OF INVENTION: Fusion Protein Genes for Treatment of  
 TITLE OF INVENTION: Aids  
 NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Kilpatrick & Cody

STREET: 100 Peachtree Street  
 CITY: Atlanta  
 STATE: Georgia  
 COUNTRY: U.S.  
 ZIP: 30303  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US92/08090  
 FILING DATE: 19920922  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pabst, Patrea L.  
 REGISTRATION NUMBER: 31,284  
 REFERENCE/DOCKET NUMBER: OMRFL29  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 404-572-6508  
 TELEFAX: 404-572-6555  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2465 base pairs  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapien  
 TISSUE TYPE: Epithelial  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: 1..6  
 OTHER INFORMATION: /note="Restriction site"  
 NAME/KEY: misc feature  
 LOCATION: 2460..2465  
 OTHER INFORMATION: /note="Restriction site"  
 PCT-US92-08090-1

Query Match 43.4%; Score 624.4; DB 5; Length 2465;  
 Best Local Similarity 99.8%; Pred. No. 6.4e-138;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGCACTGCGCTCCCTCCCA 60  
 81 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGCACTGCGCTCCCTCCCA 140  
 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAAGTACC 120  
 141 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAAGTACC 200  
 121 TGTACAGCTTCCAGAAAGAGCATACATTCCTGGAATAAATCTCCACCAATTAAG 180  
 201 TGTACAGCTTCCAGAAAGAGCATACATTCCTGGAATAAATCTCCACCAATTAAG 260  
 181 ATTCTGGGAATCAGGGCTCCTTTTAAGTCAATCAAGCTGCAAGTGAATGCGCGT 240  
 261 ATTCTGGGAATCAGGGCTCCTTTTAAGTCAATCAAGCTGCAAGTGAATGCGCGT 320  
 241 GACTCAAGAGAGGCTTTGGAGCAAGAAATTCCTCCCTGATCATCAAGATCTTAAG 300  
 321 GACTCAAGAGAGGCTTTGGAGCAAGAAATTCCTCCCTGATCATCAAGATCTTAAG 380  
 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGAGACCAAGAGAGAGGTGCAATTG 360  
 381 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGAGACCAAGAGAGAGGTGCAATTG 440  
 361 CTAGGTTCGATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 420

Db 441 CTAGTGTTCGATTGACTGCCAATCTGACACCCACCTGTCTTACGGGGCAGAGCCTGACC 500  
Qy 421 CTGACCTTGGAGAGCCCCCTGTGATAGCCCTCAGTGCATATAGAGATCCAAAGGGGT 480  
Db 501 CTGACCTTGGAGAGCCCCCTGTGATAGCCCTCAGTGCATATAGAGATCCAAAGGGGT 560  
Qy 481 AAAAACAATACAGGGGGGAGAGACCCCTCTCCGTGTCTCAGCTGAGCTCCAGATATGTGC 540  
Db 561 AAAAACAATACAGGGGGGAGAGACCCCTCTCCGTGTCTCAGCTGAGCTCCAGATATGTGC 620  
Qy 541 ACCTGACATGCACTGTCTTGCAGAAACAGAAAGGTGAGTTCAAATATGACATGTGTG 600  
Db 621 ACCTGACATGCACTGTCTTGCAGAAACAGAAAGGTGAGTTCAAATATGACATGTGTG 680  
Qy 601 GTGCTAGCTTCCAGAGGCTCCGG 626  
Db 681 GTGCTAGCTTCCAGAGGCTCCAG 706

## RESULT 8

US-08-284-391B-28  
; Sequence 28, Application US/08284391B  
; Patent No. 5851828  
; GENERAL INFORMATION:  
; APPLICANT: Seed, Brian  
; APPLICANT: Banapour, Babak  
; APPLICANT: Romeo, Charles  
; APPLICANT: Kolanus, Waldemar  
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Clark & Elbing LLP  
; STREET: 176 Federal Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/284,391B  
; FILING DATE: 02-AUG-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/195,395  
; FILING DATE: 14-FEB-1994  
; APPLICATION NUMBER: 07/847,566  
; FILING DATE: 06-MAR-1992  
; APPLICATION NUMBER: 07/665,961  
; FILING DATE: 07-MAR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Elbing, Karen L  
; REGISTRATION NUMBER: 35,238  
; REFERENCE/DOCKET NUMBER: 00786/247001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-428-0200  
; TELEFAX: 617-428-7045  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1304 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; US-08-284-391B-28

Query Match 43.2%; Score 622.8; DB 2; Length 1304;  
Best Local Similarity 99.7%; Pred. No. 1,3e-137;

Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 ATGAACCGGGAGATCCCTTTTATGACATCTTCTGGTGTGCAACTGGGCTCTCCCA 60  
Db 111 ATGAACCGGGAGATCCCTTTTATGACATCTTCTGGTGTGCAACTGGGCTCTCCCA 170  
Qy 61 GCAGCCACTCAGGAAACAAAGTGTGCTGGGCAAAAAAGGGGATACGTGAACTGACC 120  
Db 171 GCAGCCACTCAGGAAACAAAGTGTGCTGGGCAAAAAAGGGGATACGTGAACTGACC 230  
Qy 121 TGTACAGCTTCCAGAAAGACATACATTCACCTGGAATAAATCCACCATTAAG 180  
Db 231 TGTACAGCTTCCAGAAAGACATACATTCACCTGGAATAAATCCACCATTAAG 290  
Qy 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGCTGAATGATCGGCT 240  
Db 291 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGCTGAATGATCGGCT 350  
Qy 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 300  
Db 351 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 410  
Qy 301 ATGAAGACTCAGATACTTACATCTGTGAAGTGAAGACCAAGAGAGGTGCAATTG 360  
Db 411 ATGAAGACTCAGATACTTACATCTGTGAAGTGAAGACCAAGAGAGGTGCAATTG 470  
Qy 361 CTAGTGTTCGATTGACTGCAACTGTGACCACTGCTTCAAGGGGCAAGGCTGACC 420  
Db 471 CTAGTGTTCGATTGACTGCAACTGTGACCACTGCTTCAAGGGGCAAGGCTGACC 530  
Qy 421 CTGACCTTGGAGAGCCCCCTGTGATAGCCCTCAGTGCATGTAGAGATCCAAAGGGT 480  
Db 531 CTGACCTTGGAGAGCCCCCTGTGATAGCCCTCAGTGCATGTAGAGATCCAAAGGGT 590  
Qy 481 AAAAACAATACAGGGGGGAGAGACCCCTCTCCGTGTCTCAGCTGAGCTCCAGATATGTGC 540  
Db 591 AAAAACAATACAGGGGGGAGAGACCCCTCTCCGTGTCTCAGCTGAGCTCCAGATATGTGC 650  
Qy 541 ACCTGACATGCACTGTCTTGCAGAAACAGAAAGGTGAGTTCAAATATGACATGTG 600  
Db 651 ACCTGACATGCACTGTCTTGCAGAAACAGAAAGGTGAGTTCAAATATGACATGTG 710  
Qy 601 GTGCTAGCTTCCAGAGGCTCCGG 626  
Db 711 GTGCTAGCTTCCAGAGGCTCCAG 736

## RESULT 9

US-09-218-950-28  
; Sequence 28, Application US/09218950  
; Patent No. 6284240  
; GENERAL INFORMATION:  
; APPLICANT: Seed, Brian  
; APPLICANT: Banapour, Babak  
; APPLICANT: Romeo, Charles  
; APPLICANT: Kolanus, Waldemar  
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Clark & Elbing LLP  
; STREET: 176 Federal Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/218,950

FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1304 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-218-950-28

Query Match 43.2%; Score 622.8; DB 3; Length 1304;  
Best Local Similarity 99.7%; Pred. No. 1.3e-137;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTGGCACTTCTGCTGCTGCACTGGCGCTCTCCCA 60  
DB 111 ATGAACCGGGAGTCCCTTTTGGCACTTCTGCTGCTGCACTGGCGCTCTCCCA 170  
QY 61 GCAGCCACTCAGAGGAAGAGTGTGCTGGCAAAAAAGGGGATACAGTGAATGACC 120  
DB 171 GCAGCCACTCAGAGGAAGAGTGTGCTGGCAAAAAAGGGGATACAGTGAATGACC 230  
QY 121 TGTACAGCTTCCAGAGGAAGAGTGTGCTGGCAAAAAAGGGGATACAGTGAATG 180  
DB 231 TGTACAGCTTCCAGAGGAAGAGTGTGCTGGCAAAAAAGGGGATACAGTGAATG 290  
QY 181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGAATGCGCT 240  
DB 291 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGAATGCGCT 350  
QY 241 GACTCAGAAGAGCTTTGGGACCAAGAACTCCCTGATGATCAAGAATCTTAG 300  
DB 351 GACTCAGAAGAGCTTTGGGACCAAGAACTCCCTGATGATCAAGAATCTTAG 410  
QY 301 ATAGAAGACTAGATTAATCACTGTGAAGTGAAGACCAAGAGAGAGTGAATG 360  
DB 411 ATAGAAGACTAGATTAATCACTGTGAAGTGAAGACCAAGAGAGAGTGAATG 470  
QY 361 CTAGTGTTCGATTAATCACTGTGAAGTGAAGACCAAGAGAGAGTGAATG 420  
DB 471 CTAGTGTTCGATTAATCACTGTGAAGTGAAGACCAAGAGAGAGTGAATG 530  
QY 421 CTAGTGTTCGATTAATCACTGTGAAGTGAAGACCAAGAGAGAGTGAATG 480  
DB 531 CTAGTGTTCGATTAATCACTGTGAAGTGAAGACCAAGAGAGAGTGAATG 590  
QY 481 AAAAAATACAGAGGGGAGAGACCTCTCCGTGTCTCAGCTGAGCTCAGGATAGTGC 540  
DB 591 AAAAAATACAGAGGGGAGAGACCTCTCCGTGTCTCAGCTGAGCTCAGGATAGTGC 650  
QY 541 ACCTGAGATGACATGTCTTGTGCAAAACCAAGAAAGGTGAGATTCAAAATGACATG 600  
DB 651 ACCTGAGATGACATGTCTTGTGCAAAACCAAGAAAGGTGAGATTCAAAATGACATG 710  
QY 601 GTGCTAGCTTCCAGAGAGAGCTCCGG 626

DB 711 GTGCTAGCTTCCAGAGAGAGCTCCGG 736

RESULT 10  
US-08-394-388A-28  
Sequence 28, Application US/08394388A  
Patent No. 6753162  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
NUMBER OF SEQUENCES: 53  
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/394,388A  
FILING DATE: 24-FEB-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1304 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-394-388A-28

Query Match 43.2%; Score 622.8; DB 4; Length 1304;  
Best Local Similarity 99.7%; Pred. No. 1.3e-137;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTGGCACTTCTGCTGCTGCACTGGCGCTCTCCCA 60  
DB 111 ATGAACCGGGAGTCCCTTTTGGCACTTCTGCTGCTGCACTGGCGCTCTCCCA 170  
QY 61 GCAGCCACTCAGAGGAAGAGTGTGCTGGCAAAAAAGGGGATACAGTGAATGACC 120  
DB 171 GCAGCCACTCAGAGGAAGAGTGTGCTGGCAAAAAAGGGGATACAGTGAATGACC 230  
QY 121 TGTACAGCTTCCAGAGGAAGAGTGTGCTGGCAAAAAAGGGGATACAGTGAATG 180  
DB 231 TGTACAGCTTCCAGAGGAAGAGTGTGCTGGCAAAAAAGGGGATACAGTGAATG 290

```

QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGTCCATCAAGTGAATGATCGGCT 240
D 291 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGTCCATCAAGTGAATGATCGGCT 350
QY 241 GACTCAAGAAAGACCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
D 351 GACTCAAGAAAGACCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 410
QY 301 ATGAAGACTCAGTACTTACATCTGTGAAGTGGAGCAAGAGGAGGATGCAATTG 360
D 411 ATGAAGACTCAGTACTTACATCTGTGAAGTGGAGCAAGAGGAGGATGCAATTG 470
QY 361 CTAGTGTTCGATTGACTGCACTGACCACTGCTTCAAGGGGAGAGCTGACC 420
D 471 CTAGTGTTCGATTGACTGCACTGACCACTGCTTCAAGGGGAGAGCTGACC 530
QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTAGCCCTCACTGATGATTAAGATCCAGGGGT 480
D 531 CTGACCTTGGAGAGCCCCCTGTGTAGTAGCCCTCACTGATGATTAAGATCCAGGGGT 590
QY 481 AAAACATACAGGGGGGAGAGACCTCTCCGTCTGAGTGGAGCTCCAGGATAGTGC 540
D 591 AAAACATACAGGGGGGAGAGACCTCTCCGTCTGAGTGGAGCTCCAGGATAGTGC 650
QY 541 ACCTGACATGCACTGTCTTGCAGAACAGAGAGGTGAGTTCAAAATAGACATCGTG 600
D 651 ACCTGACATGCACTGTCTTGCAGAACAGAGAGGTGAGTTCAAAATAGACATCGTG 710
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
D 711 GTGCTAGCTTTCCAGAAAGGCTCCAG 736

```

## RESULT 11

```

US-08-417-495-2
; Sequence 2, Application US/08417495
; Patent No. 5843728

```

## GENERAL INFORMATION:

```

; APPLICANT: Seed, Brian et al.
; TITLE OF INVENTION: Redirection of Cellular Immunity by
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 27

```

Chimeras

```

; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston

```

```

; STATE: MA

```

```

; COUNTRY: USA

```

```

; ZIP: 02110-2804

```

## COMPUTER READABLE FORM:

```

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)

```

```

; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/417,495

```

```

; FILING DATE:

```

```

; CLASSIFICATION: 435

```

```

; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/203,866

```

```

; FILING DATE:

```

```

; APPLICATION NUMBER: US/07/847,566

```

```

; FILING DATE:

```

```

; APPLICATION NUMBER: 07/665,961

```

```

; FILING DATE: March 7, 1991

```

```

; CLASSIFICATION: 435

```

```

; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Clark, Paul T.

```

```

; REGISTRATION NUMBER: 30,162

```

```

; REFERENCE/DOCKET NUMBER: 00786/119002

```

```

; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (617) 542-5070

```

```

; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1389 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-417-495-2

```

```

Query Match 43.2%; Score 622.8; DB 2; Length 1389;
Best Local Similarity 99.7%; Pred. No. 1,3e-137;
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 ATGACCGGGAGATCCCTTTTATGACCTTGCTCTGAGTGTGCAACGAGGCTCTCCCA 60
D 1 ATGACCGGGAGATCCCTTTTATGACCTTGCTCTGAGTGTGCAACGAGGCTCTCCCA 60
QY 61 GCAGCCACTCAGGGAGAAAGAGTGTGTGCTGGCAAAAAGGGATACAGTGAACGTGACC 120
D 61 GCAGCCACTCAGGGAGAAAGAGTGTGTGCTGGCAAAAAGGGATACAGTGAACGTGACC 120
QY 121 TGTACGCTTCCAGAAAGACATACATTCACCTGGAATACTCAACCGATTAAG 180
D 121 TGTACGCTTCCAGAAAGACATACATTCACCTGGAATACTCAACCGATTAAG 180
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGTCCATCAAGTGAATGATCGGCT 240
D 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGTCCATCAAGTGAATGATCGGCT 240
QY 241 GACTCAAGAAAGACCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
D 241 GACTCAAGAAAGACCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
QY 301 ATGAAGACTCAGATACATCTGTGAAGTGGAGCAAGAGAGGTGCAATTG 360
D 301 ATGAAGACTCAGATACATCTGTGAAGTGGAGCAAGAGAGGTGCAATTG 360
QY 361 CTAGTGTTCGATTGACTGCACTGACCACTGCTTCAAGGGGAGAGCTGACC 420
D 361 CTAGTGTTCGATTGACTGCACTGACCACTGCTTCAAGGGGAGAGCTGACC 420
QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTAGCCCTCACTGATGATTAAGATCCAGGGGT 480
D 421 CTGACCTTGGAGAGCCCCCTGTGTAGTAGCCCTCACTGATGATTAAGATCCAGGGGT 480
QY 481 AAAACATACAGGGGGGAGAGACCTCTCCGTCTGAGTGGAGCTCCAGGATAGTGC 540
D 481 AAAACATACAGGGGGGAGAGACCTCTCCGTCTGAGTGGAGCTCCAGGATAGTGC 540
QY 541 ACCTGACATGCACTGTCTTGCAGAACAGAGAGGTGAGTTCAAAATAGACATCGTG 600
D 541 ACCTGACATGCACTGTCTTGCAGAACAGAGAGGTGAGTTCAAAATAGACATCGTG 600
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
D 601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626

```

## RESULT 12

```

US-08-284-391B-2
; Sequence 2, Application US/08284391B
; Patent No. 5851828

```

## GENERAL INFORMATION:

```

; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles

```

```

; ADDRESSEE: Kolanus, Waldemar

```

```

; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED

```

```

; NUMBER OF SEQUENCES: 53

```

```

; CORRESPONDENCE ADDRESS:

```



ADDRESSEE: Clark & Ebling LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391B  
FILING DATE: 02-AUG-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Eibling, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1389 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-284-391B-2

Query Match 43.2%; Score 622.8; DB 2; Length 1389;  
Best Local Similarity 99.7%; Pred. No. 1.3e-137;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTCTGTGCTGCACTGGCGCTCTCCCA 60  
1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTCTGTGCTGCACTGGCGCTCTCCCA 60  
61 GCAGCCACTCAGGAGAAAGTGTGCTGCGCAAAAAAGGGATACAGTGAACCTGACC 120  
61 GCAGCCACTCAGGAGAAAGTGTGCTGCGCAAAAAAGGGATACAGTGAACCTGACC 120  
121 TGTACAGCTTCCCAAGAGAGCATACATTCCTGAAAAAATCCCAACAGATAAAG 180  
121 TGTACAGCTTCCCAAGAGAGCATACATTCCTGAAAAAATCCCAACAGATAAAG 180  
121 TGTACAGCTTCCCAAGAGAGCATACATTCCTGAAAAAATCCCAACAGATAAAG 180  
121 TGTACAGCTTCCCAAGAGAGCATACATTCCTGAAAAAATCCCAACAGATAAAG 180  
181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCCAGCTGAATATGCGCT 240  
181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCCAGCTGAATATGCGCT 240  
181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCCAGCTGAATATGCGCT 240  
241 GACTCAAGAGAGGCTTTGGGACCAAGAAATCCCTCGATCATCAAGAACTTTAG 300  
241 GACTCAAGAGAGGCTTTGGGACCAAGAAATCCCTCGATCATCAAGAACTTTAG 300  
241 GACTCAAGAGAGGCTTTGGGACCAAGAAATCCCTCGATCATCAAGAACTTTAG 300  
301 ATAGAAGACTAGATTAATCTGTGTAGTGAAGAGAGCAAGAGAGAGGTGCAATTG 360  
301 ATAGAAGACTAGATTAATCTGTGTAGTGAAGAGAGCAAGAGAGAGGTGCAATTG 360  
301 ATAGAAGACTAGATTAATCTGTGTAGTGAAGAGAGCAAGAGAGAGGTGCAATTG 360  
361 CTAGCTTGGATTTGACTGCGCACTGTGACACCCACTGCTTCAGGGGCAAGCTGACC 420  
361 CTAGCTTGGATTTGACTGCGCACTGTGACACCCACTGCTTCAGGGGCAAGCTGACC 420  
421 CTAGCTTGGATTTGACTGCGCACTGTGACACCCACTGCTTCAGGGGCAAGCTGACC 480

421 CTGACCTTGAGAGGCCCCCTGTAGTAGCCCTCAGTGCATGTAGAGATCCAGGGGT 480  
481 AAAACATACAGGGGGGAGAACCTCTCCCTGTCTTCACTGAGCTCCAGGATGTGCG 540  
481 AAAACATACAGGGGGGAGAACCTCTCCCTGTCTTCACTGAGCTCCAGGATGTGCG 540  
541 ACCTGACATGCACTGTCTTGAGAGAACAGAGAGGAGTTCAAAATAGACATCGTG 600  
541 ACCTGACATGCACTGTCTTGAGAGAACAGAGAGGAGTTCAAAATAGACATCGTG 600  
601 GTGCTAGCTTTCAGAGAGGCTCCGG 626  
601 GTGCTAGCTTTCAGAGAGGCTCCAG 626

RESULT 13  
US-09-218-950-2  
Sequence 2, Application US/09218950  
Patent No. 6284240  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Ebling LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/218,950  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Eibling, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1389 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-09-218-950-2

Query Match 43.2%; Score 622.8; DB 3; Length 1389;  
Best Local Similarity 99.7%; Pred. No. 1.3e-137;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY      1  ATGAACCGGGGAGTCCCTTTTAGGCACTTGCTCTGATGCTGCAACTGAGGCGCTCCGCCA 60
Db      1  ATGAACCGGGGAGTCCCTTTTAGGCACTTGCTCTGATGCTGCAACTGAGGCGCTCCGCCA 60
OY      61  GCAGCACTCAGGAGAAAGAAAGTGCTGCTGGGCAAAAAAGGGGATACAGTGAATGACACC 120
Db      61  GCAGCACTCAGGAGAAACAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGAATGACACC 120
OY      121  TGTACAGCTTCCCAAGAAAGACATACAAATTCACCTGGAAAAATCTCAACCAATTAAG 180
Db      121  TGTACAGCTTCCCAAGAAAGACATACAAATTCACCTGGAAAAATCTCAACCAATTAAG 180
OY      181  ATTCTGGGAAATCAGGAGCTCTCTTTAACTAAAGTTCATCCAAAGCTGAATGATCGGCT 240
Db      181  ATTCTGGGAAATCAGGAGCTCTCTTTAAAGTTCATCCAAAGCTGAATGATCGGCT 240
OY      241  GACCTAAGAAAGAGCCTTTGGGACCAAGGAACTTCCCTGATCATCAAGAACTTTAAG 300
Db      241  GACCTAAGAAAGAGCCTTTGGGACCAAGGAACTTCCCTGATCATCAAGAACTTTAAG 300
OY      301  ATAGAAGACTCAGATATTACATCTGTGTAAGTGAAGACCAAGAGAGAGTGCATTTG 360
Db      301  ATAGAAGACTCAGATATTACATCTGTGTAAGTGAAGACCAAGAGAGAGTGCATTTG 360
OY      361  CTAGTGTTCGATTTGACTGTGCACCTCTGACACCCACCTGCTTCAAGGGGCGAGAGCTGACC 420
Db      361  CTAGTGTTCGATTTGACTGTGCACCTCTGACACCCACCTGCTTCAAGGGGCGAGAGCTGACC 420
OY      421  CTGACCTTGGAGAGCCCGCTGGTAGAGCCCTCAGTGCATGATGAAGATCCAAAGGAGT 480
Db      421  CTGACCTTGGAGAGCCCGCTGGTAGAGCCCTCAGTGCATGATGAAGATCCAAAGGAGT 480
OY      481  AAAAAATACAGGGGGGAGAAACCTCTCGTGTCTCAAGTGAAGCTCCAGATAGTGGC 540
Db      481  AAAAAATACAGGGGGGAGAAACCTCTCGTGTCTCAAGTGAAGCTCCAGATAGTGGC 540
OY      541  ACCTGACATGCACTCTTTGCAAGAACCAAGAAAGTGAAGTTCAAAAATAGCATGTG 600
Db      541  ACCTGACATGCACTCTTTGCAAGAACCAAGAAAGTGAAGTTCAAAAATAGCATGTG 600
OY      601  GTGCTAGCTTTCAGAAAGGCTCCGG 626
Db      601  GTGCTAGCTTTCAGAAAGGCTCCAG 626

RESULT 14
US-08-394-388A-2
; Sequence 2, Application US/08394388A
; Patent No. 6753162
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapur, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/394,388A
; FILING DATE: 24-FEB-1995
; CLASSIFICATION: 514

```

```

PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1389 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-394-388A-2

Query Match      43.2% Score 622.8; DB 4; Length 1389;
Beet Local Similarity 99.7%; Pred. No. 1.3e-137;
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0

QY      1 ATGAACCGGGAGTCCCTTTTAGCACTGCTTCGTGTCGCAACATGCGCTCTCCCA 60
DB      1 ATGAACCGGGAGTCCCTTTTAGCACTGCTTCGTGTCGCAACATGCGCTCTCCCA 60
QY      61 GCAGCACTCAGGGAAGAAAGATGCTGCTGGGCAAAAAAGGGAATACAGTGAACCTGACC 120
DB      61 GCAGCACTCAGGGAAGAAAGATGCTGCTGGGCAAAAAAGGGAATACAGTGAACCTGACC 120
QY      121 TGTACAGTTCCTCCAGAGAAGAGCATACATTCCTGAGAAATACTCCAACCATGAATAAG 180
DB      121 TGTACAGTTCCTCCAGAGAAGAGCATACATTCCTGAGAAATACTCCAACCATGAATAAG 180
QY      181 ATTCTGGGAATCAGGGCTCCTTTCTTAATCTAAAGTCCATCCAGCTGAATGATCGCCT 240
DB      181 ATTCTGGGAATCAGGGCTCCTTTCTTAATCTAAAGTCCATCCAGCTGAATGATCGCCT 240
QY      241 GACTCAAGAAGAGCTTTGGGACCAAGAAATCCCTCCCTGATCATCAAGATCTTAAAG 300
DB      241 GACTCAAGAAGAGCTTTGGGACCAAGAAATCCCTCCCTGATCATCAAGATCTTAAAG 300
QY      301 ATAGAAGACTCAGATCTTAATCTGTGTAAGTGAAGAGACCAAGAGAGAGGTGCAATTG 360
DB      301 ATAGAAGACTCAGATCTTAATCTGTGTAAGTGAAGAGACCAAGAGAGAGGTGCAATTG 360
QY      361 CTAGTGTTCGATTTGATCTGCAACCTGACCAACCTGCTTCAGGGGACAGAGCTTGACC 420
DB      361 CTAGTGTTCGATTTGATCTGCAACCTGACCAACCTGCTTCAGGGGACAGAGCTTGACC 420
QY      421 CTGACCTTTGAGAGGCCCCCTGGTGTGTAGCCCTCAGTGAATGATGAGAGTCCAGAGGGGT 480
DB      421 CTGACCTTTGAGAGGCCCCCTGGTGTGTAGCCCTCAGTGAATGATGAGAGTCCAGAGGGGT 480
QY      481 AAAAAACATACAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGGAACCTCAGAGTATGTGC 540
DB      481 AAAAAACATACAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGGAACCTCAGAGTATGTGC 540
QY      541 AACTGACATGCACTGTCTTTCAGAACCAAGAAAGAGTGAAGTTCAAAATGACATCTGTG 600
DB      541 AACTGACATGCACTGTCTTTCAGAACCAAGAAAGAGTGAAGTTCAAAATGACATCTGTG 600
QY      601 GTGCTAGCTTTCAGAAAGGCTCCGG 626
DB      601 GTGCTAGCTTTCAGAAAGGCTCCGG 626

```

RESULT 15  
PCT-US92-01785-2  
; Sequence 2, Application PC/TUS9201785  
; GENERAL INFORMATION:  
; APPLICANT: The General Hospital Corporation  
; TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor  
; TITLE OF INVENTION: Chimeras  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
; COMPUTER: IBM PS/2 Model 502 or 55SX  
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
; SOFTWARE: Wordperfect (Version 5.0)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/01785  
; FILING DATE: 19920306  
; CLASSIFICATION: 530  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 07/665,961  
; FILING DATE: March 7, 1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 00786/119002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ. ID NO. 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1389 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; PCT-US92-01785-2

Query Match 43.2%; Score 622.8; DB 5; Length 1389;  
Best Local Similarity 99.7%; Pred. No. 1.3e-137;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTGGCACTTGTCTGTGTCGCAACTGGCGCTCTCCCA 60  
DB 1 ATGAACCGGGAGTCCCTTTTGGCACTTGTCTGTGTCGCAACTGGCGCTCTCCCA 60  
QY 61 GCAGCCACTCAGGGGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGAACCTGACC 120  
DB 61 GCAGCCACTCAGGGGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGAACCTGACC 120  
QY 121 TGTACAGCTTCCAGAGAGAGCATATTCACACTGGAAAAAATCCCAACAGATAAAG 180  
DB 121 TGTACAGCTTCCAGAGAGAGCATATTCACACTGGAAAAAATCCCAACAGATAAAG 180  
QY 121 TGTACAGCTTCCAGAGAGAGCATATTCACACTGGAAAAAATCCCAACAGATAAAG 180  
DB 121 TGTACAGCTTCCAGAGAGAGCATATTCACACTGGAAAAAATCCCAACAGATAAAG 180  
QY 181 ATTTGGGAAATCAGGGCTCTTTTAACTAAAGTCCATCCAGGTGATATGCGCT 240  
DB 181 ATTTGGGAAATCAGGGCTCTTTTAACTAAAGTCCATCCAGGTGATATGCGCT 240  
QY 181 ATTTGGGAAATCAGGGCTCTTTTAACTAAAGTCCATCCAGGTGATATGCGCT 240  
DB 181 ATTTGGGAAATCAGGGCTCTTTTAACTAAAGTCCATCCAGGTGATATGCGCT 240  
QY 241 GACTCAGAGAGAGCTTTGGAGCAAGGAACTTCCCTGATCATCAAGATCTTAAAG 300  
DB 241 GACTCAGAGAGAGCTTTGGAGCAAGGAACTTCCCTGATCATCAAGATCTTAAAG 300  
QY 301 ATAGAAGACTCAGATATCTTACATCTGTGAAGTGAAGACAGAGAGAGGTGCAATTG 360  
DB 301 ATAGAAGACTCAGATATCTTACATCTGTGAAGTGAAGACAGAGAGAGGTGCAATTG 360

QY 361 CTAGTGTGGATTTGACTGGCCAACTGTGACACCCACCTGTTCAGGGGAGAGCTTGA 420  
DB 361 CTAGTGTGGATTTGACTGGCCAACTGTGACACCCACCTGTTCAGGGGAGAGCTTGA 420  
QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATGTAGAGATCCAGGGGT 480  
DB 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATGTAGAGATCCAGGGGT 480  
QY 481 AAAAATATACAGGGGGGAGAGACCTCTCCGTCTCAGCTGAGCTCCAGATATGAGC 540  
DB 481 AAAAATATACAGGGGGGAGAGACCTCTCCGTCTCAGCTGAGCTCCAGATATGAGC 540  
QY 541 ACCTGAGATGATCATCTGTCTTGTGAGAACAGAGAAAGGTGACTTCAAAATAGACAT 600  
DB 541 ACCTGAGATGATCATCTGTCTTGTGAGAACAGAGAAAGGTGACTTCAAAATAGACAT 600  
QY 601 GTGCTAGCTTTCAGAGAGGCTCCGG 626  
DB 601 GTGCTAGCTTTCAGAGAGGCTCCGG 626

Search completed: November 22, 2004, 14:15:34  
Job time : 143 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 12:51:30 ; Search time 774 Seconds

(without alignments)  
10053.253 Million cell updates/sec

Title: US-09-936-702-4

Perfect score: 1440

Sequence: 1 atgaccggggagtcctt.....gtcagcagtcataactcgg 1440

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3627888 seqs, 2701811610 residues

Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	625	43.4	8911	US-10-612-192-3	Sequence 3, Appli
2	624.4	43.4	1377	US-10-103-597A-38	Sequence 38, Appli
3	624.4	43.4	1377	US-10-188-444-38	Sequence 38, Appli
4	624.4	43.4	1742	US-10-151-274-7	Sequence 7, Appli
5	624.4	43.4	1742	US-10-641-643-1013	Sequence 1013, Ap
6	622.8	43.2	1389	US-09-939-537-28	Sequence 28, Appli
7	622.8	43.2	1389	US-09-939-537-2	Sequence 2, Appli
8	622.8	43.2	1389	US-09-243-008-2	Sequence 2, Appli
9	622.8	43.2	1599	US-09-939-537-3	Sequence 3, Appli
10	622.8	43.2	1599	US-09-243-008-3	Sequence 3, Appli
11	622.8	43.2	1728	US-09-939-537-1	Sequence 1, Appli
12	622.8	43.2	1728	US-09-243-008-1	Sequence 1, Appli

13	622.8	43.2	1742	US-09-891-119A-8	Sequence 8, Appli
14	622.8	43.2	1745	US-10-024-329-1	Sequence 1, Appli
15	622.8	43.2	3084	US-10-207-655-169	Sequence 169, App
16	621.2	43.1	1273	US-09-891-119A-1	Sequence 1, Appli
17	621.2	43.1	1416	US-10-157-408-2	Sequence 2, Appli
18	621.2	43.1	1416	US-10-097-044A-2	Sequence 2, Appli
19	621.2	43.1	1416	US-10-769-247-2	Sequence 2, Appli
20	613.4	42.6	1796	US-08-485-163-2	Sequence 2, Appli
21	613.4	42.6	1796	US-09-766-995-1	Sequence 1, Appli
22	612.2	42.5	1149	US-08-485-163-6	Sequence 6, Appli
23	612.2	42.5	1149	US-09-766-995-5	Sequence 5, Appli
24	612.2	42.5	3273	US-10-397-569-7	Sequence 7, Appli
25	612.2	42.5	11228	US-10-397-569-6	Sequence 6, Appli
26	612	42.5	2482	US-08-485-163-4	Sequence 4, Appli
27	612	42.5	2482	US-09-766-995-3	Sequence 3, Appli
28	609	42.3	609	US-10-466-836-24	Sequence 24, Appli
29	607.4	42.2	719	US-09-939-537-30	Sequence 30, Appli
30	571.4	39.7	573	US-10-024-329-31	Sequence 31, Appli
31	552.2	38.3	1508	US-10-157-408-5	Sequence 5, Appli
32	552.2	38.3	1508	US-10-097-044A-5	Sequence 5, Appli
33	552.2	38.3	1508	US-10-769-247-5	Sequence 5, Appli
34	549.4	38.2	1113	US-09-759-841-5	Sequence 5, Appli
35	542.4	37.7	563	US-10-073-118-24	Sequence 24, Appli
36	541.8	37.6	2379	US-10-073-118-25	Sequence 25, Appli
37	535.6	37.2	1769	US-09-934-060A-12	Sequence 12, Appli
38	535.6	37.2	2159	US-09-934-060A-1	Sequence 1, Appli
39	535.6	37.2	2159	US-09-934-060A-25	Sequence 25, Appli
40	534	37.1	534	US-09-934-060A-25	Sequence 25, Appli
41	463.2	32.2	792	US-10-422-628-13	Sequence 13, Appli
42	428.6	29.8	741	US-10-332-673-11	Sequence 11, Appli
43	408.6	28.4	723	US-10-151-882-9	Sequence 9, Appli
44	401	27.8	720	US-10-151-882-10	Sequence 10, Appli
45	401	27.8	900	US-10-203-754A-37	Sequence 37, Appli

## ALIGNMENTS

RESULT 1  
US-10-612-192-3  
; Sequence 3, Application US/10612192  
; Publication No. US20040076636A1  
GENERAL INFORMATION:  
; APPLICANT: Pal, RanaJit  
; APPLICANT: Markham, Phillip  
; APPLICANT: Keen, Timothy  
; APPLICANT: Whitney, Stephen  
; APPLICANT: Kalyanaraman, V.S.  
; TITLE OF INVENTION: HIV Immunogenic Complexes  
; FILE REFERENCE: 00711 CIP  
; CURRENT APPLICATION NUMBER: US/10/612,192  
; CURRENT FILING DATE: 2003-07-02  
; PRIOR APPLICATION NUMBER: US 09/905,962  
; PRIOR FILING DATE: 2001-07-17  
; PRIOR APPLICATION NUMBER: US 09/479,675  
; PRIOR FILING DATE: 2000-01-07  
; PRIOR APPLICATION NUMBER: US 09/075,544  
; PRIOR FILING DATE: 1998-05-11  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 8911  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Expression Vector PTK13+Neo4  
; US-10-612-192-3

Query Match 43.4%; Score 625; DB 16; Length 8911;  
Best Local Similarity 100.0%; Pred. No. 5.6e-171;  
Matches 625; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACCGGGAGGCCCTTTAGGACACTTGTCTGTGCTGCACTAGCGGCTCTCCCA 60

```
Db 3815 ATGAACGGGGAGTCCCTTTTAGGCACTTGCTTGAGTGCGCAACTGGCGCTCTCCCA 3874
Qy 61 GCAGCCACTCAGGGAAAGAAAGTGTCTGGGCAAAAAAGGGGATACAGTGAATGACC 120
Db 3875 GCAGCCACTCAGGGAAAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAATGACC 3934
Qy 121 TGTACAGCTTCCAGAAAGAGCATCAATTCCACTGSAAAAACTCCAAACAGATAAAG 180
Db 3935 TGTACAGCTTCCAGAAAGAGCATCAATTCCACTGSAAAAACTCCAAACAGATAAAG 3994
Qy 181 ATTCTGGGAATCAGGGCTCTCTTTAACTTAAAGGTCCATCAAGCTGAATGATCGGCT 240
Db 3995 ATTCTGGGAATCAGGGCTCTCTTTAACTTAAAGGTCCATCAAGCTGAATGATCGGCT 4054
Qy 241 GACTCAAGAAAGACCTTTGGGACCAAGAAACTTCCCTGATCATCAAGATCTTAAG 300
Db 4055 GACTCAAGAAAGACCTTTGGGACCAAGAAACTTCCCTGATCATCAAGATCTTAAG 4114
Qy 301 ATGAAGACTCAGATTAATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGAATTG 360
Db 4115 ATGAAGACTCAGATTAATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGAATTG 4174
Qy 361 CTAGTGTTCGGATTGACTGCACTTGACACCCACTGCTTCAAGGGCAGAGCTGACC 420
Db 4175 CTAGTGTTCGGATTGACTGCACTTGACACCCACTGCTTCAAGGGCAGAGCTGACC 4234
Qy 421 CTGACCTTGGAGAGCCCCCTGGTAGTAGCCCTCAGTGAATGAGAGTCAAGGGGT 480
Db 4235 CTGACCTTGGAGAGCCCCCTGGTAGTAGCCCTCAGTGAATGAGAGTCAAGGGGT 4294
Qy 481 AAAAACAATACAGGGGGGAGAACCTCTCCGTCTCAGCTGAGCTCCAGATAGTGGC 540
Db 4295 AAAAACAATACAGGGGGGAGAACCTCTCCGTCTCAGCTGAGCTCCAGATAGTGGC 4354
Qy 541 ACCTGACATGCACTGTCTTGCAAGACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 600
Db 4355 ACCTGACATGCACTGTCTTGCAAGACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 4414
Qy 601 GTGCTAGCTTTCCAGAAAGGCTCCG 625
Db 4415 GTGCTAGCTTTCCAGAAAGGCTCCG 4439
```

## RESULT 2

```
US-10-103-597A-38
; Sequence 38, Application US/10103597A
; Publication No. US20030096432A1
; GENERAL INFORMATION:
; APPLICANT: Jakobsen, Bent Karsten
; TITLE OF INVENTION: Screening Methods
; FILE REFERENCE: 102286.142
; CURRENT APPLICATION NUMBER: US/10/103,597A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: PCT/GB00/03579
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 992352.1
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1377)
US-10-103-597A-38
```

```
Query Match 43.4%; Score 624.4; DB 14; Length 1377;
Best Local Similarity 99.8%; Pred. No. 4.5e-171;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 ATGAACGGGGAGTCCCTTTTAGGCACTTGCTTGAGTGCGCAACTGGCGCTCTCCCA 60
Db 1 ATGAACGGGGAGTCCCTTTTAGGCACTTGCTTGAGTGCGCAACTGGCGCTCTCCCA 60
Qy 61 GCAGCCACTCAGGGAAAGAAAGTGTCTGGGCAAAAAAGGGGATACAGTGAATGACC 120
Db 61 GCAGCCACTCAGGGAAAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAATGACC 120
Qy 121 TGTACAGCTTCCAGAAAGAGCATCAATTCCACTGSAAAAACTCCAAACAGATAAAG 180
Db 121 TGTACAGCTTCCAGAAAGAGCATCAATTCCACTGSAAAAACTCCAAACAGATAAAG 180
Qy 181 ATTCTGGGAATCAGGGCTCTCTTTAACTTAAAGGTCCATCAAGCTGAATGATCGGCT 240
Db 181 ATTCTGGGAATCAGGGCTCTCTTTAACTTAAAGGTCCATCAAGCTGAATGATCGGCT 240
Qy 241 GACTCAAGAAAGACCTTTGGGACCAAGAAACTTCCCTGATCATCAAGATCTTAAG 300
Db 241 GACTCAAGAAAGACCTTTGGGACCAAGAAACTTCCCTGATCATCAAGATCTTAAG 300
Qy 301 ATGAAGACTCAGATTAATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGAATTG 360
Db 301 ATGAAGACTCAGATTAATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGAATTG 360
Qy 361 CTAGTGTTCGGATTGACTGCACTTGACACCCACTGCTTCAAGGGCAGAGCTGACC 420
Db 361 CTAGTGTTCGGATTGACTGCACTTGACACCCACTGCTTCAAGGGCAGAGCTGACC 420
Qy 421 CTGACCTTGGAGAGCCCCCTGGTAGTAGCCCTCAGTGAATGAGAGTCAAGGGGT 480
Db 421 CTGACCTTGGAGAGCCCCCTGGTAGTAGCCCTCAGTGAATGAGAGTCAAGGGGT 480
Qy 481 AAAAACAATACAGGGGGGAGAACCTCTCCGTCTCAGCTGAGCTCCAGATAGTGGC 540
Db 481 AAAAACAATACAGGGGGGAGAACCTCTCCGTCTCAGCTGAGCTCCAGATAGTGGC 540
Qy 541 ACCTGACATGCACTGTCTTGCAAGACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 600
Db 541 ACCTGACATGCACTGTCTTGCAAGACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 600
Qy 601 GTGCTAGCTTTCCAGAAAGGCTCCG 626
Db 601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626
```

## RESULT 3

```
US-10-188-444-38
; Sequence 38, Application US/10188444
; Publication No. US20030104635A1
; GENERAL INFORMATION:
; APPLICANT: Jakobsen, Bent Karsten
; TITLE OF INVENTION: Screening Methods
; FILE REFERENCE: 102286.142 (CIP)
; CURRENT APPLICATION NUMBER: US/10/188,444
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: PCT/GB00/03579
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 992352.1
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1377)
US-10-188-444-38
```

```
Query Match 43.4%; Score 624.4; DB 15; Length 1377;
Best Local Similarity 99.8%; Pred. No. 4.5e-171;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTGTGTCGACAACTGGGCTCTCCCA 60
DB 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTGTGTCGACAACTGGGCTCTCCCA 60
QY 61 GCAGCCACTCAGGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATACATGAACTGACC 120
DB 61 GCAGCCACTCAGGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATACATGAACTGACC 120
QY 121 TGTACAGCTTCCAGAAAGAAAGCATATTCCTGGAATACTCCACCAATTAAG 180
DB 121 TGTACAGCTTCCAGAAAGAAAGCATATTCCTGGAATACTCCACCAATTAAG 180
QY 181 ATTCGGGAAATCAGGGGCTCTTCTTAATAAGGTCATCCAGTGAATGATGCGCT 240
DB 181 ATTCGGGAAATCAGGGGCTCTTCTTAATAAGGTCATCCAGTGAATGATGCGCT 240
QY 241 GACTCAAGAAAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 241 GACTCAAGAAAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
QY 301 ATAGAAGACTCAGATACCTTACATCTGTGAAGTGAGAGACCAAGAGAGGATGCAATTG 360
DB 301 ATAGAAGACTCAGATACCTTACATCTGTGAAGTGAGAGACCAAGAGAGGATGCAATTG 360
QY 361 CTAGTGTGGAATGATGCTGCACTGCAACCCCTGCTTCAAGGGGCAAGGCTGACC 420
DB 361 CTAGTGTGGAATGATGCTGCACTGCAACCCCTGCTTCAAGGGGCAAGGCTGACC 420
QY 421 CTGACCTTGGAGAGGCCCCCTGTGTAGTAGCCCTCAGTGCAATGTAGAGTCCAGGGGT 480
DB 421 CTGACCTTGGAGAGGCCCCCTGTGTAGTAGCCCTCAGTGCAATGTAGAGTCCAGGGGT 480
QY 481 AAAAAATACACAGGGGGGAAAGACCTCTCTGCTGCTCAGCTGAGCTCCAGATAGTGGC 540
DB 481 AAAAAATACACAGGGGGGAAAGACCTCTCTGCTGCTCAGCTGAGCTCCAGATAGTGGC 540
QY 541 ACCTGACATGACATGCTGCTTGGAGAAACCAAGAAAGTGAATTAATACATCTGTG 600
DB 541 ACCTGACATGACATGCTGCTTGGAGAAACCAAGAAAGTGAATTAATACATCTGTG 600
QY 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626
DB 601 GTGCTAGCTTTCAGAAAGGCTCCAG 626

```

RESULT 4  
US-10-151-274-7  
Sequence 7, Application US/10151274  
Publication No. US20030064071A1

GENERAL INFORMATION:  
APPLICANT: Liltman, Dan R.  
APPLICANT: Kwon, Douglas S.  
APPLICANT: Van Kooyk, Yvette  
APPLICANT: Geltenbeck, Theo  
TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY  
TITLE OF INVENTION: INTO  
FILE REFERENCE: 1049-1-017  
CURRENT APPLICATION NUMBER: US/10/151,274  
PRIOR FILING DATE: 2002-05-20  
PRIOR APPLICATION NUMBER: US/09/517,605  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 7  
LENGTH: 1742  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-151-274-7

Query Match 43.4%; Score 624.4; DB 14; Length 1742;  
Best Local Similarity 99.8%; Pred. No. 4,9e-171;

```

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTGTGTCGACAACTGGGCTCTCCCA 60
DB 76 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTGTGTCGACAACTGGGCTCTCCCA 135
QY 61 GCAGCCACTCAGGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATACATGAACTGACC 120
DB 136 GCAGCCACTCAGGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATACATGAACTGACC 195
QY 121 TGTACAGCTTCCAGAAAGAAAGCATATTCCTGGAATACTCCACCAATTAAG 180
DB 196 TGTACAGCTTCCAGAAAGAAAGCATATTCCTGGAATACTCCACCAATTAAG 255
QY 181 ATTCGGGAAATCAGGGGCTCTTCTTAATAAGGTCATCCAGTGAATGATGCGCT 240
DB 256 ATTCGGGAAATCAGGGGCTCTTCTTAATAAGGTCATCCAGTGAATGATGCGCT 315
QY 241 GACTCAAGAAAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 316 GACTCAAGAAAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 375
QY 301 ATAGAAGACTCAGATACCTTACATCTGTGAAGTGAGAGACCAAGAGAGGATGCAATTG 360
DB 376 ATAGAAGACTCAGATACCTTACATCTGTGAAGTGAGAGACCAAGAGAGGATGCAATTG 435
QY 361 CTAGTGTGGAATGATGCTGCACTGCAACCCCTGCTTCAAGGGGCAAGGCTGACC 420
DB 436 CTAGTGTGGAATGATGCTGCACTGCAACCCCTGCTTCAAGGGGCAAGGCTGACC 495
QY 421 CTGACCTTGGAGAGGCCCCCTGTGTAGTAGCCCTCAGTGCAATGTAGAGTCCAGGGGT 480
DB 496 CTGACCTTGGAGAGGCCCCCTGTGTAGTAGCCCTCAGTGCAATGTAGAGTCCAGGGGT 555
QY 481 AAAAAATACACAGGGGGGAAAGACCTCTCTGCTGCTCAGCTGAGCTCCAGATAGTGGC 540
DB 556 AAAAAATACACAGGGGGGAAAGACCTCTCTGCTGCTCAGCTGAGCTCCAGATAGTGGC 615
QY 541 ACCTGACATGACATGCTGCTTGGAGAAACCAAGAAAGTGAATTAATACATCTGTG 600
DB 616 ACCTGACATGACATGCTGCTTGGAGAAACCAAGAAAGTGAATTAATACATCTGTG 675
QY 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626
DB 676 GTGCTAGCTTTCAGAAAGGCTCCAG 701

```

RESULT 5  
US-10-641-643-1013  
Sequence 1013, Application US/10641643  
Publication No. US20040077003A1

GENERAL INFORMATION:  
APPLICANT: Cooke, Benjamin G.  
APPLICANT: Susan G. Stuart  
APPLICANT: Jeffrey J. Selhammer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL  
GENE EXPRESSION  
NUMBER OF SEQUENCES: 1508  
CORRESPONDENCE ADDRESSES:  
ADDRESS: INCYTE PHARMACEUTICALS, INC.  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/641,643  
FILING DATE: 14-Aug-2003

CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1013:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1742 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: g179143  
SEQUENCE DESCRIPTION: SEQ ID NO: 1013 :  
US-10-641-643-1013

Query Match 43.4%; Score 624.4; DB 16; Length 1742;  
Best Local Similarity 99.8%; Pred. No. 4.9e-171;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAACGGGGAGTCCCTTTTAGGCACTTGCTGTGGTGGCAACTGGGGCTCTCCCA 60
DB 76 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGTGGTGGCAACTGGGGCTCTCCCA 135
QY 61 GCAGCCACTCAGGGAAAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAACCTGACC 120
DB 136 GCAGCCACTCAGGGAAAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAACCTGACC 195
QY 121 TGTACAGCTTCCCAAGAAAGAGCATTCATTCACCTGGAAAACTCCACCAATTAAG 180
DB 196 TGTACAGCTTCCCAAGAAAGAGCATTCATTCACCTGGAAAACTCCACCAATTAAG 255
QY 181 ATTCTGGGAATCAGGGCTCCTTTTAACATAAGGTCATCCAGCTGAATGATGGGCT 240
DB 256 ATTCTGGGAATCAGGGCTCCTTTTAACATAAGGTCATCCAGCTGAATGATGGGCT 315
QY 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 316 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 375
QY 301 ATAGAGACTCAGATCTTAATCATCTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 360
DB 376 ATAGAGACTCAGATCTTAATCATCTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 435
QY 361 CTAGTGTGGATGACTGCAACTGACACCCCACTGCTTCAAGGGGCAAGGCTGACC 420
DB 436 CTAGTGTGGATGACTGCAACTGACACCCCACTGCTTCAAGGGGCAAGGCTGACC 495
QY 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCCCTCAGTGCATGTAGAGTCCAGAGGGT 480
DB 496 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCCCTCAGTGCATGTAGAGTCCAGAGGGT 555
QY 481 AAAAACAATACAGGGGGGAAAGACCTTCTCGTGTCTCAGCTGAGCTCCAGATATGGGC 540
DB 556 AAAAACAATACAGGGGGGAAAGACCTTCTCGTGTCTCAGCTGAGCTCCAGATATGGGC 615
QY 541 ACCGTGACATGCACTGTCTTGGCAAGACCAAGAAAGGTGAGTCCAAATATGATAGTGTG 600
DB 616 ACCGTGACATGCACTGTCTTGGCAAGACCAAGAAAGGTGAGTCCAAATATGATAGTGTG 675
QY 601 GTGCTAGCTTCCAGAGGCTCCGG 626
DB 676 GTGCTAGCTTCCAGAGGCTCCAG 701
```

RESULT 6

US-09-939-537-28  
Sequence 28, Application US/09939537  
Publication No. US20030138410A1  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
Banepour, Babak  
Romeo, Charles  
Kolanue, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,537  
FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1304 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 28:  
US-09-939-537-28

Query Match 43.2%; Score 622.8; DB 10; Length 1304;  
Best Local Similarity 99.7%; Pred. No. 1.3e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGTGCTGCAACTGGGGCTCTCCCA 60
DB 111 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGTGCTGCAACTGGGGCTCTCCCA 170
QY 61 GCAGCCACTCAGGGAAAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAACCTGACC 120
DB 171 GCAGCCACTCAGGGAAAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAACCTGACC 230
QY 121 TGTACAGCTTCCCAAGAAAGAGCATATTCCTCACTGGAATACTCCACAGATTAAG 180
DB 231 TGTACAGCTTCCCAAGAAAGAGCATATTCCTCACTGGAATACTCCACAGATTAAG 290
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATCAAGGTCCATCCAGCTGAATGATCGGCT 240
DB 291 ATTCTGGGAATCAGGGCTCTTCTTAATCAAGGTCCATCCAGCTGAATGATCGGCT 350
```



```

QY 241 GACTAAGAGAGAGCCTTTGGAGCAAGGAACTTCCCTGATCATCAAGAACTCTTANG 300
DB 351 GACTAAGAGAGAGCCTTTGGAGCAAGGAACTTCCCTGATCATCAAGAACTCTTANG 410
QY 301 ATAGAAGACTCAGATTAATCACTCTGTGAAGTGAAGAGCAAGAGAGAGGTCATTTG 360
DB 411 ATAGAAGACTCAGATTAATCACTCTGTGAAGTGAAGAGCAAGAGAGAGGTCATTTG 470
QY 351 CTAGTGTTCGATTAATCACTCTGTGAAGTGAAGAGCAAGAGAGAGGTCATTTG 420
DB 471 CTAGTGTTCGATTAATCACTCTGTGAAGTGAAGAGCAAGAGAGAGGTCATTTG 530
QY 421 CTAGTGTTCGATTAATCACTCTGTGAAGTGAAGAGCAAGAGAGAGGTCATTTG 480
DB 531 CTAGTGTTCGATTAATCACTCTGTGAAGTGAAGAGCAAGAGAGAGGTCATTTG 590
QY 481 AAAAAATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 591 AAAAAATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 650
QY 541 ACCTGAGATGACATCTGTCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 651 ACCTGAGATGACATCTGTCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 710
QY 601 GTGCTAGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 626
DB 711 GTGCTAGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 736

```

# RESULT 7 US-09-939-537-2

Sequence 2, Application US/09939537  
Publication No. US20030138410A1

GENERAL INFORMATION:

APPLICANT: Seed, Brian

Banapour, Babak

Romeo, Charles

Kolans, Waldemar

TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED

CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESSEE: Clark & Elbing LLP

STREET: 176 Federal Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/939,537

FILING DATE: 24-Aug-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/284,391

FILING DATE: 02-Aug-1994

APPLICATION NUMBER: 08/195,395

FILING DATE: 14-FEB-1994

APPLICATION NUMBER: 07/847,566

FILING DATE: 06-MAR-1992

APPLICATION NUMBER: 07/665,961

FILING DATE: 07-MAR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Elbing, Karen L.

REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/247001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200

TELEFAX: 617-428-7045  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1389 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-939-537-2

Query Match 43.2%; Score 622.8; DB 10; Length 1389;  
Best Local Similarity 99.7%; Pred. No. 1.3e-170; Indels 0; Gaps 0;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGACCGGAGAGCCTTTTATGAGCACTTCTGTGCTGCACTGCGCTCTCCCA 60
DB 1 ATGACCGGAGAGCCTTTTATGAGCACTTCTGTGCTGCACTGCGCTCTCCCA 60
QY 61 GCAGCCACTCAGGGAGAAAGTGTGCTGGCAAAAAAGGGATACAGTGAATGACC 120
DB 61 GCAGCCACTCAGGGAGAAAGTGTGCTGGCAAAAAAGGGATACAGTGAATGACC 120
QY 121 TGTACAGCTTCCAGAAAGAGCATATTCACCTGGAATACTCCAAACAGATTAAG 180
DB 121 TGTACAGCTTCCAGAAAGAGCATATTCACCTGGAATACTCCAAACAGATTAAG 180
QY 181 ATTCTGGGAAATCAGAGGCTCTCTTTAACTAAAGTTCATCCAGCTGATGCGCT 240
DB 181 ATTCTGGGAAATCAGAGGCTCTCTTTAACTAAAGTTCATCCAGCTGATGCGCT 240
QY 241 GACTCAAGAGAGAGCCTTTTGGAGCAAGAGAACTTCCCTGATATCAAGATCTTAAG 300
DB 241 GACTCAAGAGAGAGCCTTTTGGAGCAAGAGAACTTCCCTGATATCAAGATCTTAAG 300
QY 301 ATAGAAGACTCAGATTAATCACTCTGTGAAGTGAAGAGCAAGAGAGAGGTCATTTG 360
DB 301 ATAGAAGACTCAGATTAATCACTCTGTGAAGTGAAGAGCAAGAGAGAGGTCATTTG 360
QY 361 CTAGTGTTCGATTAATCACTCTGTGAAGTGAAGAGCAAGAGAGAGGTCATTTG 420
DB 361 CTAGTGTTCGATTAATCACTCTGTGAAGTGAAGAGCAAGAGAGAGGTCATTTG 420
QY 421 CTAGTGTTCGATTAATCACTCTGTGAAGTGAAGAGCAAGAGAGAGGTCATTTG 480
DB 421 CTAGTGTTCGATTAATCACTCTGTGAAGTGAAGAGCAAGAGAGAGGTCATTTG 480
QY 481 AAAAAATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 481 AAAAAATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 ACCTGAGATGACATCTGTCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 541 ACCTGAGATGACATCTGTCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 601 GTGCTAGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 626
DB 601 GTGCTAGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 626

```

## RESULT 8

US-09-243-008-2  
Sequence 2, Application US/09243008  
Publication No. US2004000534A1

GENERAL INFORMATION:

APPLICANT: Seed, Brian et al.

RECEIPTOR: Receptor Chimeras

TITLE OF INVENTION: Redirection of Cellular Immunity by

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
SOFTWARE: Wordperfect (Version 5.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/243,008  
FILING DATE: 02-Feb-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/394,176  
FILING DATE: SEPTEMBER 11, 1995  
APPLICATION NUMBER: 08/203,866  
FILING DATE: February 28, 1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: March 6, 1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: March 7, 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Karen F. Lech, Ph.D  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/270001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1389 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-243-008-2

Query Match 43.2%; Score 622.8; DB 11; Length 1389;  
Best Local Similarity 99.7%; Pred. No. 1,36-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGCTGCACTGGCGCTCCTCCCA 60  
1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGCTGCACTGGCGCTCCTCCCA 60  
61 GCAGCCACTCAGGGAAGAAAGTGTCTGTGGCAAAAAGGGGATACAGTGAACCTAACC 120  
61 GCAGCCACTCAGGGAAGAAAGTGTCTGTGGCAAAAAGGGGATACAGTGAACCTAACC 120  
121 TGTAACCTTCCCGAAGAAAGAGATCAATTCACCTGGAAGAACTCCACCAATTAAG 180  
121 TGTAACCTTCCCGAAGAAAGAGATCAATTCACCTGGAAGAACTCCACCAATTAAG 180  
181 ATTCGGAATCAGGCTCCTTTTAATTAAGTCCATCCAGCTGAATGATCGGCT 240  
181 ATTCGGAATCAGGCTCCTTTTAATTAAGTCCATCCAGCTGAATGATCGGCT 240  
241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
301 ATGAAGACTCAGTACTTATCATGTGGAAGTGAAGCAAGAGGAGAGTCAATTTG 360  
301 ATGAAGACTCAGTACTTATCATGTGGAAGTGAAGCAAGAGGAGAGTCAATTTG 360  
361 CTAGTGTTCGATGACTGCAACTGCAACCCAGCTGCTTCAGGGGCAAGAGCTGACC 420  
361 CTAGTGTTCGATGACTGCAACTGCAACCCAGCTGCTTCAGGGGCAAGAGCTGACC 420  
421 CTGACCTTGAGAGAGCCCTCTGATGAGCCCTCAGTGCATGTAGAGTCCAAAGGGGT 480

421 CTGACCTTGAGAGAGCCCTCTGTAGTACCCCTCAGTGCATGTAGAGTCCAAAGGGGT 480  
481 AAAACATACAGGGGGGAGAACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGGC 540  
481 AAAACATACAGGGGGGAGAACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGGC 540  
541 ACCTGACATGACCTGTCTTTCAGAAACAGAGAAGGTGAGTTCAAAATGACATGTG 600  
541 ACCTGACATGACCTGTCTTTCAGAAACAGAGAAGGTGAGTTCAAAATGACATGTG 600  
601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626

RESULT 9  
US-09-939-537-3  
Sequence 3, Application US/09939537  
Publication No. US2003018410A1  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
Banapour, Babak  
Romeo, Charles  
Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,537  
FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1599 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-939-537-3

Query Match 43.2%; Score 622.8; DB 10; Length 1599;  
Best Local Similarity 99.7%; Pred. No. 1,4e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;



RESULT 11  
US-09-939-537-1  
; Sequence 1, Application US/09939537  
; Publication No. US20030138410A1  
; GENERAL INFORMATION:  
; APPLICANT: Seed, Brian  
; Banapur, Babak  
; Romeo, Charles  
; Kolanus, Waldemar  
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
; CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Clark & Bibling LLP  
; STREET: 176 Federal Street  
; City: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/939,537  
; FILING DATE: 24-Aug-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/284,391  
; FILING DATE: 02-AUG-1994  
; APPLICATION NUMBER: 08/195,395  
; FILING DATE: 14-FEB-1994  
; APPLICATION NUMBER: 07/847,566  
; FILING DATE: 06-MAR-1992  
; APPLICATION NUMBER: 07/665,961  
; FILING DATE: 07-MAR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Eibling, Karen L.  
; REGISTRATION NUMBER: 35,238  
; REFERENCE/DOCKET NUMBER: 00786/247001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-428-0200  
; TELEFAX: 617-428-7045  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 1:  
; LENGTH: 1728 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-939-537-1  
Query Match 43.2%; Score 622.8; DB 10; Length 1728;  
Best Local Similarity 99.7%; Pred. No. 1.4e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAACCGGAGAGTCCCTTTTAGGCACTTGTGCTGCAACTGGCGCTCTCCCA 60  
DB 1 ATGAACCGGAGAGTCCCTTTTAGGCACTTGTGCTGCAACTGGCGCTCTCCCA 60  
QY 61 GCAGCCACTAGGAGAAAGAGTGTCTGGGCAAAAAGGGGATACAGTGAATGACC 120  
DB 61 GCAGCCACTAGGAGAAAGAGTGTCTGGGCAAAAAGGGGATACAGTGAATGACC 120  
QY 121 TGTACAGCTTCCAGAGAGAGCATCAATTCCACTGSAAAAACTCCACAGATAAG 180  
DB 121 TGTACAGCTTCCAGAGAGAGCATCAATTCCACTGSAAAAACTCCACAGATAAG 180  
QY 181 ATTCTGGAAATCAGGGCTCTTCTTAACATAAGGTCATCAAGCTGAATGATGGCT 240  
DB 181 ATTCTGGAAATCAGGGCTCTTCTTAACTAAGGTCATCAAGCTGAATGATGGCT 240

DB 181 ATTCTGGAAATCAGGGCTCTTCTTAACATAAGGTCATCAAGCTGAATGATGGCT 240  
QY 241 GACTCAAGAGAGACCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 241 GACTCAAGAGAGACCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
QY 301 ATGAAGACTCAATATTCAATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTG 360  
DB 301 ATGAAGACTCAATATTCAATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTG 360  
QY 361 CTAGTGTTCGATTTAGTACTGCAACTGTGACACCACTGCTTACAGGGGAGACCTGACC 420  
DB 361 CTAGTGTTCGATTTAGTACTGCAACTGTGACACCACTGCTTACAGGGGAGACCTGACC 420  
QY 421 CTGACCTTGGAGAGCCCTCCCTGGTAGAGCCCTCAGTGAATGATGAGATCCAGGGGT 480  
DB 421 CTGACCTTGGAGAGCCCTCCCTGGTAGAGCCCTCAGTGAATGATGAGATCCAGGGGT 480  
QY 481 AAAAATACAGAGGGGGAGAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGGC 540  
DB 481 AAAAATACAGAGGGGGAGAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGGC 540  
QY 541 ACCTGACATGCACTGTCTTGGAGAACCAAGAGAGTGAATTCAAAATGACATGTG 600  
DB 541 ACCTGACATGCACTGTCTTGGAGAACCAAGAGAGTGAATTCAAAATGACATGTG 600  
QY 601 GTGCTAGCTTCCAGAGGCTCCGG 626  
DB 601 GTGCTAGCTTCCAGAGGCTCCAG 626  
RESULT 12  
US-09-243-008-1  
; Sequence 1, Application US/09243008  
; Publication No. US20040005334A1  
; GENERAL INFORMATION:  
; APPLICANT: Seed, Brian et al.  
; TITLE OF INVENTION: Redirection of Cellular Immunity by  
; Receptor Chimeras  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; City: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 50Z or 55SX  
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
; SOFTWARE: Wordperfect (Version 5.0)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/243,008  
; FILING DATE: 02-Feb-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/394,176  
; FILING DATE: SEPTEMBER 11, 1995  
; APPLICATION NUMBER: 08/203,866  
; FILING DATE: February 28, 1994  
; APPLICATION NUMBER: 07/847,566  
; FILING DATE: March 6, 1992  
; APPLICATION NUMBER: 07/665,961  
; FILING DATE: March 7, 1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Karen F. Lech, Ph.D  
; REGISTRATION NUMBER: 35,238  
; REFERENCE/DOCKET NUMBER: 00786/270001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 1:

```

? SEQUENCE CHARACTERISTICS:
?     LENGTH: 1728 base pairs
?     TYPE: nucleic acid
?     STRANDEDNESS: double
?     TOPOLOGY: linear
?     MOLECULE TYPE: DNA (genomic)
?     SEQUENCE DESCRIPTION: SEQ ID NO: 1
US-09-243-008-1

```

Query Match	43.2%	Score 622.8;	DB 11;	Length 1728;
Best Local Similarity	99.7%	Pred No. 1.4e-170;		
Matches 624;	Conservative 2;	Mismatches 0;	Indels 0;	Gaps 0

Qy	ATGACCAGGGAGTCCCTTTTATGGCACTGCTTGCGTGGTGGCACTGGAGCTCTCCCA	60
Db	1 ATGACCAGGGAGTCCCTTTTATGGCACTGCTTGCGTGGTGGCACTGGAGCTCTCCCA	60
Qy	61 GCAGCACTCAGGGAAAGAAAGTGTGTCTGGCCAAAGGAGATACGTGAACGTACC	120
Db	61 GCAGCACTCAGGGAAAGAAAGTGTGTCTGGCCAAAGGAGATACGTGAACGTACC	120
Qy	121 TGTACAGCTTCCCAAGAAAGAGCATACATTCCTGAAAAAACTCCACCAATTAAG	180
Db	121 TGTACAGCTTCCCAAGAAAGAGCATACATTCCTGAAAAAACTCCACCAATTAAG	180
Qy	181 ATTCTGGGAAATCAGGGCTCTTCTTAACTAAAGTTCATCCAGCTGAATGATCGGCT	240
Db	181 ATTCTGGGAAATCAGGGCTCTTCTTAACTAAAGTTCATCCAGCTGAATGATCGGCT	240
Qy	241 GACTCAAGAAAGCCTTTTGGGACCAAGAAACTTCCCTTGATCATCAAGATCTTAAG	300
Db	241 GACTCAAGAAAGCCTTTTGGGACCAAGAAACTTCCCTTGATCATCAAGATCTTAAG	300
Qy	301 ATGGAAGACTCAGATCTTACATCTGTGAAGTGAAGAGACCAAGAGGAAGAGTGCATTTG	360
Db	301 ATGGAAGACTCAGATCTTACATCTGTGAAGTGAAGAGACCAAGAGGAAGAGTGCATTTG	360
Qy	361 CTAGTGTGGATTTGACTGCCAATCTTGACACCACTGTCTTCAGGGCAGAGCTGACC	420
Db	361 CTAGTGTGGATTTGACTGCCAATCTTGACACCACTGTCTTCAGGGCAGAGCTGACC	420
Qy	421 CTGACCTTGGAGAGCCCCCTGTGTATGAGCCCTCAGTGCATATGTAAGAGTCCAGGGGT	480
Db	421 CTGACCTTGGAGAGCCCCCTGTGTATGAGCCCTCAGTGCATATGTAAGAGTCCAGGGGT	480
Qy	481 AAAAACAATACAGGGGAGGAAGACCTCTCCGTCTCAGCTGAGAGCTCAGAGTATGAGGC	540
Db	481 AAAAACAATACAGGGGAGGAAGACCTCTCCGTCTCAGCTGAGAGCTCAGAGTATGAGGC	540
Qy	541 ACCTGACATGCACTGTCTTTGCAAGACCAAGAAAGGTGAGATTCAAAATGACATCGTG	600
Db	541 ACCTGACATGCACTGTCTTTGCAAGACCAAGAAAGGTGAGATTCAAAATGACATCGTG	600
Qy	601 GTGCTAGCTTTCCAGAAAGGCTTCCGG 626	
Db	601 GTGCTAGCTTTCCAGAAAGGCTTCCAG 626	

```

1  RESULT 13
2  US-09-891-119A-8
3  Sequence 8, Application US/09891119A
4  Publication No. US20040013683A1
5  GENERAL INFORMATION:
6  APPLICANT: Maddon, Paul J.
7  TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4
8  FILE REFERENCE: 24577-CT-B
9  CURRENT APPLICATION NUMBER: US/09/891,119A
10 CURRENT FILING DATE: 2001-06-25
11 NUMBER OF SEQ ID NOS: 22
12 SOFTWARE: PatentIn version 3.1
13 SEQ ID NO 8
14 LENGTH: 1742
15 TYPE: DNA
16

```

ORGANISM: Human  
US-09-891-119A-8

Query Match	43.2%;	Score 622.8;	DB 11;	Length 1742;
Best Local Similarity	99.7%;	Pred. No. 1.4e-170;		
Matches 624;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0

Oy	1	ATTGAACCGGGGAGTCCCTTTTAAAGCATTGCTCTGTGGTCTCAACTGGCGCTCTCCCA	60
Db	76	ATTGAACCGGGGAGTCCCTTTTAAAGCATTGCTCTGTGGTCTCAACTGGCGCTCTCCCA	135
Oy	61	GCAGCCACTCAGGGGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATPACAGTGAACCTGACC	120
Db	136	GCAGCCACTCAGGGGAAACAAAGTGTGCTGGGCAAAAAAGGGGATPACAGTGAACCTGACC	195
Oy	121	TGTACAGCTTCCCGAAGAAAGGCACTAATATTCACCTGGAAAAATCCCAACCAAGATAAG	180
Db	196	TGTACAGCTTCCCGAAGAAAGGCACTAATATTCACCTGGAAAAATCCCAACCAAGATAAG	255
Oy	181	ATTTCGGAAATCAGGGGCTCCTCTTAACTAAAGGTCCATCCAAAGTGAATGATCGCGCT	240
Db	256	ATTTCGGAAATCAGGGGCTCCTCTTAACTAAAGGTCCATCCAAAGTGAATGATCGCGCT	315
Oy	241	GACTCAAGAAAGAGCGCTTTGGGACCAAGAAATCTTCCCTGATCATCAAGATCTTAAAG	300
Db	316	GACTCAAGAAAGAGCGCTTTGGGACCAAGAAATCTTCCCTGATCATCAAGATCTTAAAG	375
Oy	301	ATTGAAGACTCAGATTACTTAATCATCTGTGAAGTGGAGGACACAGAAGGAGGAGCAATTG	360
Db	376	ATTGAAGACTCAGATTACTTAATCATCTGTGAAGTGGAGGACACAGAAGGAGGAGCAATTG	435
Oy	361	CTAGTGTCCGATTGACTGCGCAACTCTGACACCCACTTGTTCAGGGGACAGGCTTGACC	420
Db	436	CTAGTGTCCGATTGACTGCGCAACTCTGACACCCACTTGTTCAGGGGACAGGCTTGACC	495
Oy	421	CTGACCTTTGGAGAGCGCCCTGTGTATGTAGCCCTCAAGTGCAATGTATGAGTCCAAAGGGT	480
Db	496	CTGACCTTTGGAGAGCGCCCTGTGTATGTAGCCCTCAAGTGCAATGTATGAGTCCAAAGGGT	555
Oy	481	AAAAACATACAGGGGGGAGAGACCTCTCCGTGTCTCAGCTGGAGCTCCAGGATAGTGGC	540
Db	556	AAAAACATACAGGGGGGAGAGACCTCTCCGTGTCTCAGCTGGAGCTCCAGGATAGTGGC	615
Oy	541	ACCTGGAATGCACTGTCTTTGCAAGACCAAGAAAGGTGGAAGTTCAAAATGACATCGTG	600
Db	616	ACCTGGAATGCACTGTCTTTGCAAGACCAAGAAAGGTGGAAGTTCAAAATGACATCGTG	675
Oy	601	GTCGTAGCTTCCAGAAAGGCGCTCGG 626	
Db	676	GTCGTAGCTTCCAGAAAGGCGCTCCAG 701	

```

RESULT 14
US-10-024-329-1
; Sequence 1, Application US/10024329
; Publication No. US20030157063A1
; GENERAL INFORMATION:
; APPLICANT: SAMRADI, Kamel
; APPLICANT: TOURAINE, Jean-Louis
; APPLICANT: LEROY, Pierre
; APPLICANT: MERTALI, Majid
; TITLE OF INVENTION: Gene therapy using anti-GP41 antibody and cd4 immunoadhesin
; FILE REFERENCE: 109993
; CURRENT APPLICATION NUMBER: US/10/024,329
; CURRENT FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1745
; TYPE: DNA
; ORGANISM: human
US-10-024-329-1

```

Query Match 43.2%; Score 622.8; DB 15; Length 1745;  
 Best Local Similarity 99.7%; Pred. No. 1.4e-170;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTCTGTGTGCAACTGGCGCTCTCCCA 60
DB 79 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGTGCAACTGGCGCTCTCCCA 138
QY 61 GCAGCCACTCAGGAGAAAGTGTGTGTGGCAAAAAAGGGGATACAGTGAATGACC 120
DB 139 GCAGCCACTCAGGAGAAAGTGTGTGTGGCAAAAAAGGGGATACAGTGAATGACC 198
QY 121 TGTACAGCTTCCAGAAAGAGCATACATTTCCACTGGAATACTCCAACAGATAAG 180
DB 199 TGTACAGCTTCCAGAAAGAGCATACATTTCCACTGGAATACTCCAACAGATAAG 258
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACATAAGTCCATCCAGCTGAATGATCGGCT 240
DB 259 ATTCTGGGAATCAGGGCTCTTCTTAACATAAGTCCATCCAGCTGAATGATCGGCT 318
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 300
DB 319 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 378
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGAATG 360
DB 379 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGAATG 438
QY 361 CTAGTGTTCGATTTGACTGCCAATCTGACACCACTGCTTCAAGGGGAGAGCTTAAC 420
DB 439 CTAGTGTTCGATTTGACTGCCAATCTGACACCACTGCTTCAAGGGGAGAGCTTAAC 498
QY 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCCCTCAGTGAATGAGAGTCCAAGGGGT 480
DB 499 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCCCTCAGTGAATGAGAGTCCAAGGGGT 558
QY 481 AAAAACAATACAGGGGGGAGAGACCTCTCGTGTCTGAGCTGAGAGTCCAAGATAGTGC 540
DB 559 AAAAACAATACAGGGGGGAGAGACCTCTCGTGTCTGAGCTGAGAGTCCAAGATAGTGC 618
QY 541 ACCTGACATGACATGCTGTTTGCAGAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 600
DB 619 ACCTGACATGACATGCTGTTTGCAGAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 678
QY 601 GTGCTAGCTTCCAGAGGCTCCGG 626
DB 679 GTGCTAGCTTCCAGAGGCTCCAG 704

```

## RESULT 15

US-10-207-655-169  
 ; Sequence 169, Application US/10207655  
 ; Publication No. US20030118592A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ledbetter, Jeffrey A.  
 ; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS  
 ; FILE REFERENCE: 380069.401C1  
 ; CURRENT APPLICATION NUMBER: US/10/207,655  
 ; CURRENT FILING DATE: 2002-07-25  
 ; NUMBER OF SEQ ID NOS: 426  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 169  
 ; LENGTH: 3084  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-207-655-169

Query Match 43.2%; Score 622.8; DB 15; Length 3084;  
 Best Local Similarity 99.7%; Pred. No. 1.7e-170;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTCTGTGTGCAACTGGCGCTCTCCCA 60

```

DB 153 ATGAACCGGGAGTCCCTTTTAGGCACTTGTCTGTGTGCAACTGGCGCTCTCCCA 212
QY 61 GCAGCCACTCAGGAGAAAGTGTGTGTGGCAAAAAAGGGGATACAGTGAATGACC 120
DB 213 GCAGCCACTCAGGAGAAAGTGTGTGTGGCAAAAAAGGGGATACAGTGAATGACC 272
QY 121 TGTACAGCTTCCAGAAAGAGCATACATTTCCACTGGAATACTCCAACAGATAAG 180
DB 273 TGTACAGCTTCCAGAAAGAGCATACATTTCCACTGGAATACTCCAACAGATAAG 332
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACATAAGTCCATCCAGCTGAATGATCGGCT 240
DB 333 ATTCTGGGAATCAGGGCTCTTCTTAACATAAGTCCATCCAGCTGAATGATCGGCT 392
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 300
DB 393 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 452
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGAATG 360
DB 453 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGAATG 512
QY 361 CTAGTGTTCGATTTGACTGCCAATCTGACACCACTGCTTCAAGGGGAGAGCTTAAC 420
DB 513 CTAGTGTTCGATTTGACTGCCAATCTGACACCACTGCTTCAAGGGGAGAGCTTAAC 572
QY 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCCCTCAGTGAATGAGAGTCCAAGGGGT 480
DB 573 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCCCTCAGTGAATGAGAGTCCAAGGGGT 632
QY 481 AAAAACAATACAGGGGGGAGAGACCTCTCGTGTCTGAGCTGAGAGTCCAAGATAGTGC 540
DB 633 AAAAACAATACAGGGGGGAGAGACCTCTCGTGTCTGAGCTGAGAGTCCAAGATAGTGC 600
QY 541 ACCTGACATGACATGCTGTTTGCAGAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 600
DB 693 ACCTGACATGACATGCTGTTTGCAGAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 752
QY 601 GTGCTAGCTTCCAGAGGCTCCGG 626
DB 753 GTGCTAGCTTCCAGAGGCTCCAG 778

```

Search completed: November 22, 2004, 16:11.13  
 Job time : 778 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 09:24:30 ; Search time 4897 Seconds  
(without alignments)  
10715.380 Million cell updates/sec

Title: US-09-936-702-4

Perfect score: 1440

Sequence: 1 atgacacggggagctccctt.....gtcagcagtcataactg 1440

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3282875 seqs, 1821965908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_g881:\*  
9: gb\_g882:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	622.8	43.2	783	6	CD609192 56069931J
2	622.8	43.2	932	5	BX438745 BX438745
3	610.2	42.4	816	4	BI838409 603083373
4	599.4	41.6	822	5	BX457159 BX457159
5	598.8	41.6	801	4	BI822118 603039813
6	589	40.9	840	4	BI838893 603087279
7	577.4	40.1	791	4	BI819727 603041415
8	539	37.4	713	5	BX437619 BX437619
9	519.4	36.1	723	4	BI915265 603184313
10	506.4	35.2	679	6	CD609191 56069931H
11	504	35.0	732	1	AUI41298 AUI41298
12	488.8	33.9	754	7	CO246446 AGENCOURT
13	488.4	33.9	710	6	CF125444 UI-HF-EL0
14	483.8	33.6	694	7	CB052686 NISC_g109
15	480.6	33.4	803	6	CO249540 AGENCOURT
16	466	32.4	809	7	CO248894 AGENCOURT
17	457.4	31.8	748	5	BX378811 BX378811
18	452.8	31.4	821	7	CO245985 AGENCOURT
19	429.4	28.8	616	6	CF125140 UI-HF-EL0
20	407.6	28.0	751	7	CO248544 AGENCOURT
21	402.6	26.0	687	2	AW752367 RC2-CT020
22	374.6	25.5	767	1	AUI36822 AUI36822
23	366.6	25.5	686	6	CB052685 NISC_g109
24	364.2	25.3	755	7	CO246525 AGENCOURT

C 25	354.8	24.6	638	2	AW752358 RC2-CT020
26	345	24.0	893	1	AL553528 AL553528
27	328.8	22.8	707	7	CK835626 CK835626
28	302.8	21.0	567	6	CB997693 AGENCOURT
29	261.2	18.1	336	2	AM407282 UI-HF-BL0
30	261.2	18.1	374	2	AM406380 UI-HF-BL0
31	261.2	18.1	733	4	BG685053 602636816
32	259.6	18.0	325	2	AM407857 UI-HF-BL0
33	259.6	18.0	409	2	AM406057 UI-HF-BL0
34	259.6	18.0	429	6	CD707673 EST24200
35	259.6	18.0	448	2	AM404726 UI-HF-BL0
36	259.6	18.0	499	6	CD703936 EST20463
37	259.6	18.0	514	2	AM405207 UI-HF-BL0
38	259.6	18.0	643	4	BM782582 K-EST0059
39	259.6	18.0	723	4	BM007674 603617103
40	259.6	18.0	1047	5	BU146881 AGENCOURT
41	258	17.9	330	2	AM406212 UI-HF-BL0
42	258	17.9	414	6	CD693211 EST9734 h
43	258	17.9	425	2	AM406785 UI-HF-BL0
44	258	17.9	465	2	AM404697 UI-HF-BL0
45	258	17.9	471	6	CD687139 CD687139

## ALIGNMENTS

RESULT 1  
LOCUS CD609192/c 783 bp mRNA linear EST 12-JAN-2004  
DEFINITION 56069931J1 FLP Homo sapiens cDNA, mRNA sequence.  
ACCESSION CD609192  
VERSION CD609192.1 GI:40257455  
KEYWORDS EST.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L. 1 (bases 1 to 783)  
TITLE Circular rapid amplification of cDNA ends for high-throughput extension cloning of partial genes  
JOURNAL Genomics 84 (1), 205-210 (2004)  
COMMENT Contact: Fu GK  
Incyte Genomics, Inc.  
3160 Porter Dr., Palo Alto, CA 94304, USA  
Tel: 6508454102  
Email: gfu@incyte.com.

FEATURES  
source location/Qualifiers  
1..783  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="FLP"  
/note="Vector: pdrive Cloning Vector"

ORIGIN  
Query Match 43.2%; Score 622.8; DB 6; Length 783;  
Best Local Similarity 99.7%; Pred. No. 2.4e-152;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAACGGGGAGTCCCTTTTAGGCACTTCTGCTGCTGCAACTGGCGCTCTCCCA 60  
|||  
26 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGCTGCTGCAACTGGCGCTCTCCCA 667  
|||  
61 GCAGCCACTCCAGGAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
|||  
66 GCAGCCACTCCAGGAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 607  
|||  
121 TGTACAGCTTCCAGAGAAAGCATACATTCACCTGGAATACTCCACGATTAAG 180  
|||  
606 TGTACAGCTTCCAGAGAAAGCATACATTCACCTGGAATACTCCACGATTAAG 547  
|||  
181 ATTGGGAATCAGGGCTCTTCTTACTTAAGATCCATCCAGCTGAATGATGGCGCT 240

```

|||||
Db 546 ATCTGGGAATCAGGGCTCCTTCTTAACATAAGCTCATCAAGTGAATGATCGGCT 487
Qy 241 GACTCAAGAAAGACCTTTGGGACCAAGAAACCTTCCCGGATCATCAAGATCTTAAG 300
Db 486 GACTCAAGAAAGACCTTTGGGACCAAGAAACCTTCCCGGATCATCAAGATCTTAAG 427
Qy 301 ATGAAAGACTCAGATTAATCATCTGTGAAGTGAGGACCAAGAGAGAGTCAATTG 360
Db 426 ATGAAAGACTCAGATTAATCATCTGTGAAGTGAGGACCAAGAGAGAGTCAATTG 367
Qy 361 CTAGTGTTCGATTAATCATCTGTGAAGTGAGGACCAAGAGAGAGTCAATTG 420
Db 366 CTAGTGTTCGATTAATCATCTGTGAAGTGAGGACCAAGAGAGAGTCAATTG 307
Qy 421 CTGACCTTGGAGAGCCCCCTGTAGAGCCCTCAGTGCAATGAGAGTCAAGGGGT 480
Db 306 CTGACCTTGGAGAGCCCCCTGTAGAGCCCTCAGTGCAATGAGAGTCAAGGGGT 247
Qy 481 AAAAACAATACAGGGGGGGAAGACCTCTCGTGTCTGAGCTGAGCTCAAGATAGTGC 540
Db 246 AAAAACAATACAGGGGGGGAAGACCTCTCGTGTCTGAGCTGAGCTCAAGATAGTGC 187
Qy 541 ACCTGACATGACATGCTGCTTGCAGAACCAAGAAAGTGAAGTTCAAATAGATGCTG 600
Db 186 ACCTGACATGACATGCTGCTTGCAGAACCAAGAAAGTGAAGTTCAAATAGATGCTG 127
Qy 601 GTGCTAGCTTCCAGAAAGGCTCCGG 626
Db 126 GTGCTAGCTTCCAGAAAGGCTCCAG 101

```

```

RESULT 2
LOCUS BX438745 932 bp mRNA linear EST 05-MAY-2004
DEFINITION BX438745 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE002YG01
5-PRIME, mRNA sequence.
ACCESSION BX438745
VERSION BX438745.2 GI:47035651
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 932)
L.J.W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30787731.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 6485.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdnats=CS0DE002AD01QP1&c=6485.r.
Location/Qualifiers
1..932
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE002YG01"
/rnause_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector."

```

## FEATURES

## SOURCE

```

1..932
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE002YG01"
/rnause_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector."

```

Library was not normalized."

Query Match 43.2%; Score 622.8; DB 5; Length 932;  
 Best Local Similarity 99.7%; Pred. No. 2; se-152;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 ATGAAACGGGAGATCCCTTTTAAGCACTTCTCTGTGTGTGCAACTGCGCTCTCCCA 60
Db 180 ATGAAACGGGAGATCCCTTTTAAGCACTTCTCTGTGTGTGCAACTGCGCTCTCCCA 239
Qy 61 GCAGCCACTCAGGGGAAAGAAAGTGTGCTGGGCAAAAAGGGATACATGAACTGACC 120
Db 240 GCAGCCACTCAGGGGAAAGAAAGTGTGCTGGGCAAAAAGGGATACATGAACTGACC 299
Qy 121 TGTACACCTTCCAGAAAGAGATGATCAATTCACCTGSAAAAACCTCAACGATTAAG 180
Db 300 TGTACACCTTCCAGAAAGAGATGATCAATTCACCTGSAAAAACCTCAACGATTAAG 359
Qy 181 ATCTGGGAATACAGGGCTCCTTTTAAGTAAAGTTCATCAAGCTGAATGATCGGCT 240
Db 360 ATCTGGGAATACAGGGCTCCTTTTAAGTAAAGTTCATCAAGCTGAATGATCGGCT 419
Qy 241 GACTCAAGAAAGACCTTTGGGACCAAGAAACCTTCCCGTATCATCAAGATCTTAAG 300
Db 420 GACTCAAGAAAGACCTTTGGGACCAAGAAACCTTCCCGTATCATCAAGATCTTAAG 479
Qy 301 ATGAAAGACTCAGATTAATCATCTGTGAAGTGAGGACCAAGAGAGAGTGAATTG 360
Db 480 ATGAAAGACTCAGATTAATCATCTGTGAAGTGAGGACCAAGAGAGAGTGAATTG 539
Qy 361 CTAGTGTTCGATTAATCATCTGTGAAGTGAGGACCAAGAGAGAGTGAATTG 420
Db 540 CTAGTGTTCGATTAATCATCTGTGAAGTGAGGACCAAGAGAGAGTGAATTG 599
Qy 421 CTGACCTTGGAGAGCCCCCTGTAGAGCCCTCAGTGCAATGAGAGTCCAAAGGGGT 480
Db 600 CTGACCTTGGAGAGCCCCCTGTAGAGCCCTCAGTGCAATGAGAGTCCAAAGGGGT 659
Qy 481 AAAAACAATACAGGGGGGGAAGACCTCTCGTGTCTGAGCTGAGCTCCAGATAGTGC 540
Db 660 AAAAACAATACAGGGGGGGAAGACCTCTCGTGTCTGAGCTGAGCTCCAGATAGTGC 719
Qy 541 ACCTGACATGACATGCTGCTTGCAGAACCAAGAAAGTGAAGTTCAAATAGATGCTG 600
Db 720 ACCTGACATGACATGCTGCTTGCAGAACCAAGAAAGTGAAGTTCAAATAGATGCTG 779
Qy 601 GTGCTAGCTTCCAGAAAGGCTCCGG 626
Db 780 GTGCTAGCTTCCAGAAAGGCTCCAG 805

```

```

RESULT 3
LOCUS BI838409 816 bp mRNA linear EST 04-OCT-2001
DEFINITION BI838409
603063373F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5222323 5',
mRNA sequence.
ACCESSION BI838409.1 GI:15949959
VERSION BI838409
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 816)
NIH-MGC http://mgc.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)

```

```

REFERENCE
AUTHORS NIH-MGC
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)

```



DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/BLML ac:  
<http://image.llnl.gov>  
 Plate: LLM11559 row: b column: 20  
 High quality sequence stop: 803.  
 Location/Qualifiers

## FEATURES

source

1. 816  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:522323"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_120"  
 /note="Organ: pooled pancreas and spleen; Vector:  
 pCMV-Sport6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
 source anonymous pool of spleen and pancreas from 28 yo  
 male. Library is oligo-dT primed and directionally cloned  
 (EcoRV site is destroyed upon cloning). Average insert  
 size 1.5 kb, insert size range 1-2.5 kb. Library is  
 normalized and enriched for full-length clones and was  
 constructed by C. Gruber (Invitrogen). Research Genetics  
 tracking code 025. Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 42.4%; Score 610.2; DB 4; Length 816;  
 Best Local Similarity 99.4%; Pred. No. 5e-149;  
 Matches 623; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

Oy 1 ATGAACCGGGAGTCCCTTTTATGACCTTGTGTCGCTCAACGTGGCGCTCCCTCCA 60
Db 167 ATGAACCGGGAGTCCCTTTTATGACCTTGTGTCGCTCAACGTGGCGCTCCCTCCA 226
Oy 61 GCAGCACTCAGGAGAAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACGACC 120
Db 227 GCAGCACTCAGGAGAAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACGACC 286
Oy 121 TGTACAGCTTCCCAAGAAAGCATATCTCCAGTGAAGAACTCCACCAAGATTAAG 180
Db 287 TGTACAGCTTCCCAAGAAAGCATATCTCCAGTGAAGAACTCCACCAAGATTAAG 346
Oy 181 ATTCTGGGAATCAGGGGCTCTTCTTAATGAAGTCCATCCAGCTGAATGATGCGCT 240
Db 347 ATTCTGGGAATCAGGGGCTCTTCTTAATGAAGTCCATCCAGCTGAATGATGCGCT 406
Oy 241 GACTCAGAAGAGCCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 300
Db 407 GACTCAGAAGAGCCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 466
Oy 301 ATAGAAGCTCAGATCTTCACTCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 360
Db 467 ATAGAAGCTCAGATCTTCACTCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 526
Oy 361 CTAGGTTCCGATTAAGTCCAACTCTGACACCACTGCTTCAAGGGGAGAGCTGACC 420
Db 527 CTAGGTTCCGATTAAGTCCAACTCTGACACCACTGCTTCAAGGGGAGAGCTGACC 586
Oy 421 CTGACCTTGAAGAGCCCTCTGTAGTACCCCTCAGTGCATGTAGAGTCCAGAGGAT 480
Db 587 CTGACCTTGAAGAGCCCTCTGTAGTACCCCTCAGTGCATGTAGAGTCCAGAGGAT 646
Oy 481 AAAAAATACAGAGGGGGGAAGACCTCTTC -GTGTCTCAGCTGAGCTCCAGGATAGTG 539
Db 647 AAAAAATACAGAGGGGGGAAGACCTCTTC -GTGTCTCAGCTGAGCTCCAGGATAGTG 706
Oy 540 CACCTGACATGCACTGTCTTGACAGACCAAGAAAGGTGAGTCAAAATGACATGT 599
Db 707 CACCTGACATGCACTGTCTTGACAGACCAAGAAAGGTGAGTCAAAATGACATGT 766
Oy 600 GGTGCTAGCTTTCCAGAGGCTCTCCG 626
Db 767 GGTGCTAGCTTTCCAGAGGCTCTCCAG 793

```

## RESULT 4

BX457159

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BX457159 822 bp mRNA linear EST 05-MAY-2004  
 BX457159 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP006YP07  
 5-PRIME, mRNA sequence.  
 BX457159  
 BX457159.2 GI:47051693  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 822)  
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On May 22, 2003 this sequence version replaced gi:31020883.  
 CONTACT: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 Division of Invitrogen.  
 This sequence belongs to sequence cluster 6485. r  
 For more information about this cluster, see  
<http://www.genoscope.cns.fr/cdnasr-CS0CAP006CH04QP1&c=6485.r>.

## FEATURES

source

1. 822  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0CAP006YP07"  
 /tissue\_type="THYMUS"  
 /clone\_lib="Homo sapiens THYMUS"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo(dT) primer. Five prime end enriched,  
 double-strand cDNA was digested with Not I and cloned into  
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
 Library was not normalized."

## ORIGIN

Query Match 41.6%; Score 599.4; DB 5; Length 822;  
 Best Local Similarity 99.2%; Pred. No. 3.4e-146;  
 Matches 622; Conservative 1; Mismatches 2; Indels 2; Gaps 2;

```

Oy 1 ATGAACCGGGAGTCCCTTTTATGACCTTGTGTCGCTCAACGTGGCGCTCCCTCCA 60
Db 183 ATGAACCGGGAGTCCCTTTTATGACCTTGTGTCGCTCAACGTGGCGCTCCCTCCA 242
Oy 61 G-CAGCCACTCAGGAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACGACC 119
Db 243 G-CAGCCACTCAGGAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACGACC 302
Oy 120 CTGTACAGCTTCCCAAGAAAGCATATCTCCAGTGAAGAACTCCACCAAGATTAAG 179
Db 303 CTGTACAGCTTCCCAAGAAAGCATATCTCCAGTGAAGAACTCCACCAAGATTAAG 362
Oy 180 GATTCTGGGAATACAGGGGCTCTTCTTAATGAAGTCCATCCAGCTGAATGATGCGC 239
Db 363 GATTCTGGGAATACAGGGGCTCTTCTTAATGAAGTCCATCCAGCTGAATGATGCGC 422
Oy 240 TGACCTCAAGAGAGCCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTA 299
Db 423 TGACCTCAAGAGAGCCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTA 482
Oy 300 GATAGAAGACTCAGATCTTCACTCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATT 359
Db 483 GATAGAAGACTCAGATCTTCACTCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATT 542

```

QY 360 GCTAGTGTGGATTGATGCTCCAACTCTGACACCCACCTGCTTCAGGGGAGAGCCTGAC 419  
 Db 543 GCTAGTGTGGATTGATGCTCCAACTCTGACACCCACCTGCTTCAGGGGAGAGCCTGAC 602  
 QY 420 CCTGACCTTGGAGAGCCCCCTGCTGATAGCCCTCTGATGCAATGTAGAGTCCAAAGGG 479  
 Db 603 CCGACCTTGGAGAGCCCCCTGCTGATAGCCCTCTGATGCAATGTAGAGTCCAAAGGG 662  
 QY 480 TAAAAACATACAGGGGGGAGAGACCTCTCGTGTCTCAGTGGAGCTCCAGATAGTGG 539  
 Db 663 TAAAAACATACAGGGGGGAGAGACCTCTCGTGTCTCAGTGGAGCTCCAGATAGTGG 721  
 QY 540 CACCTGGACATGACCTGTCTTGGAGAACCAAGAGAGTGGAGTTCAAAATAGACATCGT 599  
 Db 722 CACCTGGACATGACCTGTCTTGGAGAACCAAGAGAGTGGAGTTCAAAATAGACATCGT 781  
 QY 600 GGTGCTAGCTTTCCAGAGGCTCTCCG 626  
 Db 782 GGTGCTAGCTTTCCAGAGGCTCTCCG 808

RESULT 5  
 B1822118 801 bp mRNA linear EST 04-OCT-2001  
 LOCUS 603039813F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5180642 5',  
 DEFINITION mRNA sequence.

ACCESSION B1822118  
 VERSION B1822118  
 KEYWORDS B1822118.1 GI:15933668  
 SOURCE EST.  
 ORGANISM Homo sapiens (human)

REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 TITLE 1 (bases 1 to 801)  
 JOURNAL NIH-MGC http://mgc.nci.nih.gov/  
 COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-r@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: LLM1450 row: j column: 03  
 High quality sequence stop: 793.  
 Location/Qualifiers

FEATURES  
 source

1. 801  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5180642"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH MGC 115"  
 /note="Organ: pooled brain, lung, testis; Vector:  
 pCMV-SPORT6; Site 1: NotI; site2: EcoRV (destroyed); RNA  
 source anonymous pool of 6 male brains, age range 23-27; 1  
 male lung, age 27; and 1 male testis, age 69. Library is  
 oligo-dT primed and directionally cloned (EcoRV site is  
 destroyed upon cloning). Average insert size 1.8 kb,  
 insert size range 1-3 kb. Library is normalized and  
 enriched for full-length clones and was constructed by C.  
 Gruber (Invitrogen). Research Genetics tracking code  
 021. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 41.6%; Score 598.8; DB 4; Length 801;  
 Best Local Similarity 99.4%; Pred. No. 4.9e-146;  
 Matches 622; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 ATGAACGGGAGATCCCTTTAGGCACTTCTGTGTCGCAACTGGCGCTCTCCCA 60

|||||  
 Db 175 ATGAACGGGAGATCCCTTTAGGCACTTCTGTGTCGCAACTGGCGCTCTCCCA 234  
 QY 61 GCAGCACTCAGGAGAAAGTGTGCTGGGCAAAAAGGGGATCAGTGAACCTGACC 120  
 Db 235 GCAGCACTCAGGAGAAAGTGTGCTGGGCAAAAAGGGGATCAGTGAACCTGACC 294  
 QY 121 TGTACAGCTTCCGAGAAAGACATATCAATTCCTCAGTGAACCAACGATTAAG 180  
 Db 295 TGTACAGCTTCCGAGAAAGACATATCAATTCCTCAGTGAACCAACGATTAAG 354  
 QY 181 ATTCTGGGAATCAGGCTCTCTTTTAACTAAAGTTCATCCAACTGATGATCGGCT 240  
 Db 355 ATTCTGGGAATCAGGCTCTCTTTTAACTAAAGTTCATCCAACTGATGATCGGCT 414  
 QY 241 GACTCAAGAAAGACCTTTGGGACCAAGGAACTCCCTGATCATCAAGATCTTAAG 300  
 Db 415 GACTCAAGAAAGACCTTTGGGACCAAGGAACTTTCCCTGATCATCAAGATCTTAAG 474  
 QY 301 ATGAAGACTCAGATATCTTACATCTGTGAAGTGAAGGACCAAGAGAGTGAATTG 360  
 Db 475 ATGAAGACTCAGATATCTTACATCTGTGAAGTGAAGGACCAAGAGAGTGAATTG 534  
 QY 361 CTAGTGTTCGATTTGACTGCAACTGTGACACCACTGCTTCAAGGCGCAGAGCTGACC 420  
 Db 535 CTAGTGTTCGATTTGACTGCAACTGTGACACCACTGCTTCAAGGCGCAGAGCTGACC 594  
 QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGAAGTGAAGTGAAGTGAAGGCT 480  
 Db 595 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGAAGTGAAGTGAAGTGAAGGCT 654  
 QY 481 AAAAACATACA-GGGGGGAGAGACCTCTCTCGTGTCTCAGCTGAGCTCCAGATAGTGG 539  
 Db 655 AAAAACATACAGGGGGGAGAGACCTCTCTCGTGTCTCAGCTGAGCTCCAGATAGTGG 714  
 QY 540 CACCTGGACATGACCTGTCTTGGAGAACCAAGAGAGTGAAGTCAAAATAGACATCGT 599  
 Db 715 CACCTGGACATGACCTGTCTTGGAGAACCAAGAGAGTGAAGTCAAAATAGACATCGT 774  
 QY 600 -GGTGTAGCTTCCAGAGGCTCC 624  
 Db 775 GGTGTAGCTTCCAGAGGCTCC 800

RESULT 6  
 B1838893 840 bp mRNA linear EST 04-OCT-2001  
 LOCUS 603087279F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5226427 5',  
 DEFINITION mRNA sequence.

ACCESSION B1838893  
 VERSION B1838893.1 GI:15950443  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 TITLE 1 (bases 1 to 840)  
 JOURNAL NIH-MGC http://mgc.nci.nih.gov/  
 COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-r@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: LLM1569 row: m column: 20  
 High quality sequence start: 11  
 High quality sequence stop: 832.  
 Location/Qualifiers

FEATURES

source

1. .840  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5226427"  
/lab\_host="DH10B"  
/note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 40.9%; Score 589; DB 4; Length 840;  
Best Local Similarity 98.1%; Pred. No. 1.9e-143;  
Matches 617; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTCTGTCTGCACTGCGCTCTCCCA 60  
168 ATGAACCGGGAGTCCCTTTTAGGCACTTGTCTGTCTGCACTGCGCTCTCCCA 227  
61 GGAGCCACTCAGGAAAGAAAGTGTGTGTGGGCAAAAAGGGATACAGTGAATGAC 120  
228 GGAGCCACTCAGGAAAGAAAGTGTGTGTGGGCAAAAAGGGATACAGTGAATGAC 287  
121 TGTACAGTCTCCAGAAAGAGCATATCAATTCATGAAAAATCCAAACGATTAAG 180  
288 TGTACAGTCTCCAGAAAGAGCATATTCATGAAAAATCCAAACGATTAAG 347  
181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTCCATCCAGCTGATGTCGCT 240  
348 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTCCATCCAGCTGATGTCGCT 407  
241 GACTCAAGAAAGCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 300  
408 GACTCAAGAAAGCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 467  
301 ATAGAAGACTGATATCAATCTGTGAAGTGAAGAGACCAAGAGAGAGGCAATTG 360  
468 ATAGAAGACTGATATCAATCTGTGAAGTGAAGAGACCAAGAGAGAGGCAATTG 527  
361 CTAGTGTTCGATGACTGCAACTCTGACCACTGCTTCAGGGGACAGGCTGACC 420  
528 CTAGTGTTCGATGACTGCAACTCTGACCACTGCTTCAGGGGACAGGCTGACC 587  
421 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCTCATGTCATGTAAGAGTCCAGGGGT 480  
588 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCTCATGTCATGTAAGAGTCCAGGGGT 647  
481 AAA-AACATACAGGGGGGAGAACCTCTCGTGTCTCAGCTGAGAGCTCAAGATATG 539  
648 AAAACAATACAGGGGGGAGAACCTCTCGTGTCTCAGCTGAGAGCTCAAGATATG 707  
540 CACCTGACATGACATGTCTTCAGAAACCAAGAAAGAGTGAATCAAAATAGACATCGT 599  
708 CACCTGACATGACATGTCTTCAGAAACCAAGAAAGAGTGAATCAAAATAGACATCGT 767  
600 GG-TGCTAGCTTCCAGAAAGGCTCCGGA 627  
768 GGAGGCTAGCTTCCAGAAAGGCTCCGGA 796

RESULT 7  
BI819727 791 bp mRNA linear EST 04-OCT-2001  
LOCUS BI819727 791bp mRNA linear EST 04-OCT-2001  
DEFINITION 603041415p1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5182274 5',  
ACCESSION mRNA sequence.  
BI819727

VERSION BI819727.1 GI:15931277  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE NIH-MGC <http://mgc.mci.nih.gov/>,  
National Institutes of Health, Mammalian Gene Collection (MGC)  
1 (bases 1 to 791)  
AUTHORS NIH-MGC  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: LLM11454 row: n column: 03  
High quality sequence stop: 782.

FEATURES

source

1. .791  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5182274"  
/lab\_host="DH10B"  
/clone\_1ib="NIH MGC 115"  
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 40.1%; Score 577.4; DB 4; Length 791;  
Best Local Similarity 99.3%; Pred. No. 2.1e-140;  
Matches 611; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTCTGTCTGCACTGCGCTCTCCCA 60  
176 ATGAACCGGGAGTCCCTTTTAGGCACTTGTCTGTCTGCACTGCGCTCTCCCA 235  
61 GGAGCCACTCAGGAAAGAAAGTGTGTGTGGGCAAAAAGGGATACAGTGAATGAC 120  
236 GGAGCCACTCAGGAAAGAAAGTGTGTGTGGGCAAAAAGGGATACAGTGAATGAC 295  
121 TGTACAGTCTCCAGAAAGAGCATATTCATGAAAAATCCAAACGATTA-A 179  
236 TGTACAGTCTCCAGAAAGAGCATATTCATGAAAAATCCAAACGATTA-A 355  
180 GATTCGGAATATCAGGGCTCTTCTTAATTAAGTCCATCAAGCTGAATGATGCGCG 239  
236 TGTACAGTCTCCAGAAAGAGCATATTCATGAAAAATCCAAACGATTA-A 355  
356 GATTCGGAATATCAGGGCTCTTCTTAATTAAGTCCATCAAGCTGAATGATGCGCG 415  
240 TGACTCAAGAAAGAGCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTA 299  
416 TGACTCAAGAAAGAGCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTA 475  
300 GATAGAAGACTCAGATCTTACATCTGAGAGTGAAGAGACCAAGAGAGAGGTCAATT 359  
476 GATAGAAGACTCAGATCTTACATCTGAGAGTGAAGAGACCAAGAGAGAGGTCAATT 535  
360 GCTAGTGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 419  
536 GCTAGTGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 595

Qy 420 CCTGACCTTGG-AGAGCCCCCTGTGTAGTACCCCTCAGTGCATGTAGAGTCCAAAGG 478  
 Db 596 CTTGACCTTGGCAAGCCCCCTGTGTAGTACCCCTCAGTGCATGTAGAGTCCAAAGG 655  
 Qy 479 GTAAAAATATCAGGGGGGAAAGACCTTCCGTGTCTCAGCTGAGAGTCCAGATATG 538  
 Db 656 GTAAAAATATCAGGGGGGAAAGACCTTCCGTGTCTCAGCTGAGAGTCCAGATATG 715  
 Qy 539 GCACCTGAGAC-TGCATGTCTTGCAGAACCAAGAAAGGTGGAGTCAAAATAGACATC 597  
 Db 716 GCACCTGAGACATTCATGTCTTGCAGAACCAAGAAAGGTGGAGTCAAAATAGACATC 775  
 Qy 598 GTGTGCTAGCTTTC 612  
 Db 776 GTGTGCTAGCTTTC 790

RESULT 8  
 BX437619 713 bp mRNA linear EST 04-MAY-2004  
 LOCUS BX437619 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP007YD02  
 DEFINITION 5-PRIME, mRNA sequence.  
 ACCESSION BX437619  
 VERSION BX437619.2 GI:46999918  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 713)  
 AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT On May 15, 2003 this sequence version replaced gi:30771588.  
 CONTACT: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.  
 This sequence belongs to sequence cluster 6485.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdnas=CS0CAP007YD01Q1&c=6485.r.  
 FEATURES  
 source  
 1..713  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0CAP007YD02"  
 /clone\_type="THYMUS"  
 /clone\_lib="Homo sapiens THYMUS"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN  
 Query Match 37.4%; Score 539; DB 5; Length 713;  
 Best Local Similarity 99.5%; Pred. No. 2.5e-110;  
 Matches 550; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 1 ATGAACGGGGAGTCCCTTTAGCACTTGTCTGTGTGCAACTGGCGCTCTCCCA 60  
 Db 161 ATGAACGGGGAGTCCCTTTAGCACTTGTCTGTGTGCAACTGGCGCTCTCCCA 220  
 Qy 61 GCAGCCACTCAGGGAAGAAAGTGTGTGTGGCAAAAAAGGGATACAGTGAATGACC 120  
 Db 221 GCAGCCACTCAGGGAAGAAAGTGTGTGTGGCAAAAAAGGGATACAGTGAATGACC 280

Qy 121 TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGGAAGAAAACTCCAAACGATTAAG 180  
 Db 281 TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGGAAGAAAACTCCAAACGATTAAG 340  
 Qy 181 ATTCTGGGAATCAGGGCTCTCTTTACTTAAGGTTCATCCAAAGTGAATGATCGGCT 240  
 Db 341 ATTCTGGGAATCAGGGCTCTCTTTACTTAAGGTTCATCCAAAGTGAATGATCGGCT 400  
 Qy 241 GACTCAAGAAAGAGCTTTGGGACCAAGGAAAGTCCCGCTGATCATCAAGAAATCTTAAG 300  
 Db 401 GACTCAAGAAAGAGCTTTGGGACCAAGGAAAGTCCCGCTGATCATCAAGAAATCTTAAG 460  
 Qy 301 ATAGAACTCAGATTAATCTATCTGTGAAGTGGAGCAAGAAAGAGAGTGCATTTG 360  
 Db 461 ATAGAACTCAGATTAATCTATCTGTGAAGTGGAGCAAGAAAGAGAGTGCATTTG 520  
 Qy 361 CTAGTGTTCGAGATTGACTGCAACTGTGACACCCACTGCTTCAGGGGCAAGCTTACC 420  
 Db 521 CTAGTGTTCGAGATTGACTGCAACTGTGACACCCACTGCTTCAGGGGCAAGCTTACC 580  
 Qy 421 CTGACCTTGGAGAGCCCCCTGGTAGAGCCCTCAGTGCATGTAGAGTCCAAAGGGGT 480  
 Db 581 CTGACCTTGGAGAGCCCCCTGGTAGAGCCCTCAGTGCATGTAGAGTCCAAAGGGGT 640  
 Qy 481 AAAAATATCA-GGGGGGGAAAGACCTCTCCGTGTCTCAGCTGAGAGTCCAGATTAAG 539  
 Db 641 AAAAATATCA-GGGGGGGAAAGACCTCTCCGTGTCTCAGCTGAGAGTCCAGATTAAG 700  
 Qy 540 CACCTGAGCATGC 552  
 Db 701 CACCTGAGCATGC 713

RESULT 9  
 BI915265 723 bp mRNA linear EST 16-OCT-2001  
 LOCUS BI915265 60318431F1 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5248226 5'  
 DEFINITION mRNA sequence.  
 ACCESSION BI915265  
 VERSION BI915265.1 GI:16179363  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 723)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strauberg, Ph.D.  
 Email: sgabbs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LNL11626 row: J column: 03  
 High quality sequence stop: 717.  
 FEATURES  
 source  
 1..723  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5248226"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH MGC 121"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon



Fax: 81-438-52-3986  
 Email: genomics@hri.co.jp  
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix  
 Research Institute; cDNA library construction: Department of  
 Virology, Institute of Medical Science, University of Tokyo, and  
 Helix Research Institute.

# FEATURES

Source  
 Location/Qualifiers  
 1. 732  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="THYR0100359"  
 /tissue="thyroid gland"  
 /clone\_id="THYR01"  
 /note="Vector: pME18SFLJ"

## ORIGIN

Query Match 35.0%; Score 504; DB 1; Length 732;  
 Best Local Similarity 97.2%; Pred. No. 4e-121;  
 Matches 522; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

```

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTGCTGCTGCTGCACTGGGCTCTCCCA 60
   |||||
Db 196 ATGAACCGGGAGTCCCTTTAGGCACTGCTGCTGCTGCACTGGGCTCTCCCA 255
   |||||
QY 61 GCAGCCACTAGGAGAAAGTGTGCTGGCAAAAAGGGATAGAGTGAATGACC 120
   |||||
Db 256 GCAGCCACTAGGAGAAAGTGTGCTGGCAAAAAGGGATAGAGTGAATGACC 315
   |||||
QY 121 TGTACAGCTTCCGAAAGAGCATACATTCACCTGGAATACTCCACAGATTAAG 180
   |||||
Db 316 TGTACAGCTTCCGAAAGAGCATACATTCACCTGGAATACTCCACAGATTAAG 375
   |||||
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACCTAAAGTCCATCAAGTGAATGATCGGCT 240
   |||||
Db 376 ATTCTGGGAATCAGGGCTCTTCTTAACCTAAAGTCCATCAAGTGAATGATCGGCT 435
   |||||
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
   |||||
Db 436 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 495
   |||||
QY 301 ATGAAGACTCAGATTAATCTTATCTGTGAAGTGAAGCAAGAGAGAGTGAATTG 360
   |||||
Db 496 ATGAAGACTCAGATTAATCTTATCTGTGAAGTGAAGCAAGAGAGAGTGAATTG 555
   |||||
QY 361 CTAGTGTTCGATTAAGTCACTGCACTGCAACCCACCTGCTTCAAGGGGCAAGCTGACC 420
   |||||
Db 556 CTAGTGTTCGATTAAGTCACTGCACTGCAACCCACCTGCTTCAAGGGGCAAGCTGACC 615
   |||||
QY 421 CTGACCTTGGAGAGCCCCCTGTGATAGACCCCTCAAGTCAATGAGAGTCAAGGGGT 480
   |||||
Db 616 CTGACCTTGGAGAGCCCCCTGTGATAGACCCCTCAAGTCAATGAGAGTCAAGGGGT 675
   |||||
QY 481 AAAAACAATACAGGGGGG--AAGACCTCTCTCGTCTCAAGTGAAGCTCCAGATA 535
   |||||
Db 676 AAAAACAATACAGGGGGG--AAGACCTCTCTCGTCTCAAGTGAAGCTCCAGATA 732
   |||||

```

RESULT 12  
 CO246446 754 bp mRNA linear EST 23-JUN-2004

LOCUS  
 DEFINITION  
 AGENCOURT 26517764 NIH MGC 212 Homo sapiens cDNA clone  
 IMAGE:30924065 5', mRNA sequence.

ACCESSION  
 CO246446  
 CO246446.1 GI:49111174

VERSION  
 CO246446.1

KEYWORDS  
 EST.

ORGANISM  
 Homo sapiens (human)

SOURCE  
 Homo sapiens

REFERENCE  
 1 (bases 1 to 754)

AUTHORS  
 NIH-MGC http://mgc.ncl.nih.gov/.

TITLE  
 National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)  
 Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: egadsb-f@mail.nih.gov

Tissue Procurement: Mary Hendrix  
 cDNA Library Preparation: M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov

Plate: NDAM165 row: 0 column: 18  
 High quality sequence stop: 507.

## FEATURES

### source

1. 754  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30924065"  
 /tissue\_type="Chondrosarcoma lung Metastasis cell lines"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_id="NIH\_MGC\_212"  
 /note="Organ: Lung; Vector: pYX-Asc; Site 1: Ecor I;  
 Site 2: Not I; The library was constructed according  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction. Ligated  
 with Ecor I adaptor, digested with Not I and then cloned  
 directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is GATTAAGCCCA. Tissue was provided by Mary Hendrix."

## ORIGIN

Query Match 33.9%; Score 488.8; DB 7; Length 754;  
 Best Local Similarity 96.6%; Pred. No. 4.1e-117;  
 Matches 510; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

```

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTCTGCTGCTGCACTGGGCTCTCCCA 60
   |||||
Db 173 ATGAACCGGGAGTCCCTTTAGGCACTTCTGCTGCTGCACTGGGCTCTCCCA 232
   |||||
QY 61 GCAGCCACTAGGAGAAAGTGTGCTGGCAAAAAGGGATAGAGTGAATGACC 120
   |||||
Db 233 GCAGCCACTAGGAGAAAGTGTGCTGGCAAAAAGGGATAGAGTGAATGACC 292
   |||||
QY 121 TGTACAGCTTCCGAAAGAGCATACATTCACCTGGAATACTCCACAGATTAAG 180
   |||||
Db 293 TGTACAGCTTCCGAAAGAGCATACATTCACCTGGAATACTCCACAGATTAAG 352
   |||||
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACCTAAAGTCCATCAAGTGAATGATCGGCT 240
   |||||
Db 353 ATTCTGGGAATCAGGGCTCTTCTTAACCTAAAGTCCATCAAGTGAATGATCGGCT 412
   |||||
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
   |||||
Db 413 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 472
   |||||
QY 301 ATGAAGACTCAGATTAATCTTATCTGTGAAGTGAAGCAAGAGAGAGTGAATTG 360
   |||||
Db 473 ATGAAGACTCAGATTAATCTTATCTGTGAAGTGAAGCAAGAGAGAGTGAATTG 532
   |||||
QY 361 CTAGTGTTCGATTAAGTCACTGCACTGCAACCCACCTGCTTCAAGGGGCAAGCTGACC 420
   |||||
Db 533 CTAGTGTTCGATTAAGTCACTGCACTGCAACCCACCTGCTTCAAGGGGCAAGCTGACC 592
   |||||
QY 421 CTGACCTTGGAGAGCCCCCTGTGATAGACCCCTCAAGTCAATGAGAGTCAAGGGGT 480
   |||||
Db 593 CTGACCTTGGAGAGCCCCCTGTGATAGACCCCTCAAGTCAATGAGAGTCAAGGGGT 652
   |||||

```

QY 481 AAAACATACAGGGGGAGAACCCCTCTCGTGTCTACGTGAGCTC 528  
 DB 653 AAAA-ATACATGTAGGGGAAAACCTCTCTGTCTCATCTGATCTC 699

RESULT 13  
 CFI25444 710 bp mRNA linear EST 05-AUG-2003  
 LOCUS IMAGE:30563734 5', mRNA sequence.  
 DEFINITION CFI25444  
 CFI25444.1 GI:33201679  
 EST.  
 VERSION CFI25444  
 KEYWORDS Homo sapiens (human)  
 SOURCE  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 710)  
 AUTHORS Ronaldo, M.F., Lennon, G. and Soares, M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8889548  
 COMMENT Contact: Soares, MB  
 Coordinated laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Mary Hendrix  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
 http://genome.uiowa.edu/distribution/humanfl.html  
 Seq primer: PYX-5

FEATURES  
 source location/Qualifiers  
 1..710  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30563734"  
 /tissue\_type="Chondrosarcoma Lung Metastasis cell lines"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_1lb="NIH MGC 212"  
 /note="Organ: Lung; Vector: pYX-Asc; Site 1: EcoR I;  
 Site 2: Not I; The library was constructed according  
 to Bonaldi, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated  
 with EcoR I adaptor, digested with Not I and then cloned  
 directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is GATTAAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN  
 Query Match 33.9%; Score 488.4; DB 6; Length 710;  
 Best Local Similarity 98.9%; Pred No. 5,1e-117;  
 Matches 522; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGTCTGCAACGTGGCGCTCTCCCA 60  
 DB 186 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGTCTGCAACGTGGCGCTCTCCCA 245  
 QY 61 GCAGCACTCTCAGGGAAGAAAGTGTCTGTGGGCAAAAAAGGAGATACAGTGAATGACC 120  
 DB 246 GCAGCACTCTCAGGGAAGAAAGTGTGTGTGGGCAAAAAAGGAGATACAGTGAATGACC 305  
 QY 121 GTTACAGCTTCCAGAGAAGACATATTCACCTGGAAGAAAGTCCACACAGATAAG 180

DB 306 TGTACAGCTTCCAGAGAAGACATATTCACCTGGAAGAAAGTCCACAGATAAG 365  
 QY 181 ATTCTGGGAATACAGGCTCTCTTTAACTAAAGGTCAATCCAGCTGAATATGCGCT 240  
 DB 366 ATTCTGGGAATACAGGCTCTCTTTAACTAAAGGTCAATCCAGCTGAATATGCGCT 425  
 QY 241 GACTGAAGAAGAGCTTTGGGACCAAGAACTTCCCTGATATCAAGATCTTAAAG 300  
 DB 426 GACTGAAGAAGAGCTTTGGGACCAAGAACTTCCCTGATATCAAGATCTTAAAG 485  
 QY 301 ATAGAAGACTCAGATCTTACATCTGTGAAGTGAGAGACCAAGAGAGAGGTGAATTG 360  
 DB 486 ATAGAAGACTCAGATCTTACATCTGTGAAGTGAGAGACCAAGAGAGAGGTGAATTG 545  
 QY 361 CTAGTGTTCGATGACTCCCACTGTGACACCCAGCTGTTCAAGGGCAGAGCTGACC 420  
 DB 546 CTAGTGTTCGATGACTCCCACTGTGACACCCAGCTGTTCAAGGGCAGAGCTGACC 605  
 QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTGATGCAATGTAGAGTCAAGGGGT 480  
 DB 606 CTGACCTTGGAGAG-CCCCCTGTGTAGTAC-CCCTGATGCAATGTAGAGTCAAGGGGT 663  
 QY 481 AAAACATACAGGGGGAGAACCCCTCTCGTGTCTACGTGAGCTC 528  
 DB 664 AAAA-ATACATGTAGGGGAAAACCTCTCTGTCTCATCTGATCTC 710

RESULT 14  
 CB052686 694 bp mRNA linear EST 17-JAN-2003  
 LOCUS NISC\_g109606.y1 NCI\_CGAP\_Le12 Homo sapiens cDNA clone IMAGE:3290578  
 DEFINITION 5', mRNA sequence.  
 ACCESSION CB052686  
 VERSION CB052686.1 GI:27790973  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 694)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-rcmail.nih.gov  
 cDNA Library Preparation:  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium/LNL  
 Sequencing Center (NISC)  
 Clone Distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 info@image.lnl.gov  
 Plate: LLM8058 row: I column: 11  
 Seq primer: M13RP1 reverse primer (ABI).  
 Location/Qualifiers  
 1..694  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3290578"  
 /tissue\_type="leiomyosarcoma"  
 /lab\_host="DH10B"  
 /clone\_1lb="NCI CGAP Le12"  
 /note="Organ: soft tissue; Vector: pRTD-Pac (Pharmacia)  
 with a modified polylinker; Site 1: Not I; Site 2: Eco RI;  
 1st strand cDNA was primed with a Not I - oligo(dT) primer  
 (5'-ACTGGAAGATTTCGCGCCGACATCTTTTCTTTTCTTTT-3'),  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pRTD vector. Library  
 went through one round of normalization. Library



ORIGIN constructed by Bento Soares and M. Fatima Bonaldo."

Query Match 33.4%; Score 483.8; DB 6; Length 694;  
 Best Local Similarity 99.4%; Pred. No. 8.2e-116;  
 Matches 496; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGCTGCTGCAACTGCGCTCTCCCA 60  
 DB 196 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGCTGCTGCAACTGCGCTCTCCCA 255

QY 61 GCAGCCACTCAGGGAAGAAAGTGTGTGTCGCGC-AAAAAGGGATACAGTGAATGAC 119  
 DB 256 GCAGCCACTCAGGGAAGAAAGTGTGTGTCGCGCAAAAAAGGGATACAGTGAATGAC 315

QY 120 CTGTAGAGCTTCCGAGAAAGAGCATCAATTCACCTGGAATACTCCAAACGATPAA 179  
 DB 316 CTGTAGAGCTTCCGAGAAAGAGCATCAATTCACCTGGAATACTCCAAACGATPAA 375

QY 180 GATTCTGGGAAATCAGGGCTCTCTTAACTAAAGTCCATCAAGCTGAATGATCGCG 239  
 DB 376 GATTCTGGGAAATCAGGGCTCTCTTAACTAAAGTCCATCAAGCTGAATGATCGCG 435

QY 240 TGACTCAAGAAAGCCCTTTGGACCAAGAAATTCCTCCCTGATCATCAAGAAATCTTAA 299  
 DB 436 TGACTCAAGAAAGCCCTTTGGACCAAGAAATTCCTCCCTGATCATCAAGAAATCTTAA 495

QY 300 GATGAAGACTCAGATCACTTCACTGTGTGAAGTGAAGACCAAGAAAGAGTGCATTT 359  
 DB 496 GATGAAGACTCAGATCACTTCACTGTGTGAAGTGAAGACCAAGAAAGAGTGCATTT 555

QY 360 GCTAGTTCGGATTTGATGCTGCAACTCTGACACCCACCTGCTTCAAGGGGAGAGCTGAC 419  
 DB 556 GCTAGTTCGGATTTGATGCTGCAACTCTGACACCCACCTGCTTCAAGGGGAGAGCTGAC 615

QY 420 CTGTAGCTTGGAGAGCCCTCTGTAGTACCCCTCACTGATGATGAGAGTCCAGGGG 479  
 DB 616 CTGTAGCTTGGAGAGCCCTCTGTAGTACCCCTCACTGATGATGAGAGTCCAGGGG 675

QY 480 TAAAAACATACAGGGGGGG 498  
 DB 676 TAAAAACATACAGGGGGGG 694

RESULT 15  
 CO249540 803 bp mRNA linear EST 23-JUN-2004  
 LOCUS DEFINITION  
 AGENCOURT\_26527156 NIH\_MGC\_212 Homo sapiens cDNA clone  
 IMAGE:30921364 5', mRNA sequence.  
 ACCESSION CO249540  
 VERSION CO249540.1 GI:49120610  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 803)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: gcapbs-r@mail.nih.gov  
 Tissue Procurement: Mary Hendrix  
 cDNA Library Preparation: M. Bento Soares, University of Iowa  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: NDAM158 row: 0 column: 05

FEATURES High quality sequence stop: 525.  
 Location/Qualifiers  
 1..803

source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30921364"  
 /tissue\_type="Chondrosarcoma Lung Metastasis cell lines"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_1b="NIH MGC 212"  
 /note="Organ: Lung; Vector: pYX-Asc; Site: 1: Ecor I;  
 Site: 2: Not I; The library was constructed according  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated  
 with Ecor I adaptor, digested with Not I and then cloned  
 directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is GATTAAGCCA. Tissue was provided by Mary Hendrix."

## ORIGIN

Query Match 33.4%; Score 480.6; DB 7; Length 803;  
 Best Local Similarity 96.5%; Pred. No. 5.9e-115;  
 Matches 501; Conservative 0; Mismatches 16; Indels 2; Gaps 1;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGCTGCTGCAACTGCGCTCTCCCA 60  
 DB 173 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGCTGCTGCAACTGCGCTCTCCCA 232

QY 61 GCAGCCACTCAGGGAAGAAAGTGTGTGTCGCGCAAAAAAGGGATACAGTGAATGAC 120  
 DB 233 GCAGCCACTCAGGGAAGAAAGTGTGTGTCGCGCAAAAAAGGGATACAGTGAATGAC 232

QY 121 TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGGAATACTCCAAACGATPAA 180  
 DB 293 TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGGAATACTCCAAACGATPAA 352

QY 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGCTGAATGATCGGCT 240  
 DB 353 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGCTGAATGATCGGCT 412

QY 241 GACTCAAGAAAGCCCTTTGGGACCAAGAAATTCCTCCCTGATCATCAAGAAATCTTAA 300  
 DB 413 GACTCAAGAAAGCCCTTTGGGACCAAGAAATTCCTCCCTGATCATCAAGAAATCTTAA 472

QY 301 ATGAAGACTCAGATCACTTCACTGTGTGAAGTGAAGACCAAGAAAGAGTGCATTTG 360  
 DB 473 ATGAAGACTCAGATCACTTCACTGTGTGAAGTGAAGACCAAGAAAGAGTGCATTTG 532

QY 361 CTAGTGTGGATTTGATGCTGCAACTGACACCCACCTGTTAGGGGGAGAGCTGAGC 420  
 DB 533 CTAGTGTGGATTTGATGCTGCAACTGACACCCACCTGTTAGGGGGAGAGCTGAGC 592

QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGAATGATGATGATGATGATG 480  
 DB 593 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGAATGATGATGATGATGATG 650

QY 481 AAAAATACAGGGGGGAGAGACCTCTCCGTGTTCA 519  
 DB 651 TAAAAATACAGGGGGGAGAGACCTCTCCGTGTTCA 689

Search completed: November 22, 2004, 14:13:11  
 Job time : 4905 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:29:16 ; Search time 55.6204 Seconds

(without alignments)  
878.780 Million cell updates/sec

Title: US-09-936-702-3

Perfect score: 2668

Sequence: 1 MNRGVPRFHLVLVQLALLP.....GTRLEIKVPRSGHHHHH 508

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR79:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1064	39.9	458	1	RWHUT4
2	916	34.3	432	1	RMWCT4
3	831	31.1	432	1	RMWQT4
4	750	28.1	268	2	A56446
5	668	25.0	249	2	S41374
6	618.5	22.2	459	2	M46254
7	616	22.1	233	2	UC5322
8	596	22.3	432	2	S30193
9	508.5	19.1	457	2	A27449
10	507	19.0	128	2	S40343
11	498	18.7	107	2	S34005
12	498	18.7	117	2	S40362
13	498	18.7	457	1	RMWST4
14	485.5	18.2	116	2	B25555
15	485	18.2	123	2	S40378
16	481.5	18.0	131	2	S40328
17	480.5	18.0	109	1	K3HUPM
18	475.5	17.8	215	2	UB0244
19	466	17.5	111	2	S23628
20	460	17.2	108	2	G44151
21	458	17.2	128	2	A56701
22	456.5	17.1	114	2	S54905
23	456	17.1	128	2	S40345
24	455	17.1	114	2	S46375
25	455	17.1	129	2	S23627
26	454	17.0	115	1	K3HUC8
27	453	17.0	98	2	I30608
28	453	17.0	115	2	A30553
29	453	17.0	144	2	PL0106

30	451	16.9	95	2	PH0868	Ig kappa chain V r
31	451	16.9	215	2	JR0243	Ig kappa chain NIG
32	449	16.8	110	2	S40326	Ig kappa chain V-J
33	448.5	16.8	109	2	H30601	Ig kappa chain V-I
34	448	16.8	128	2	S40379	Ig kappa chain V-J
35	445.5	16.7	627	2	S14683	Ig mu chain precu
36	443.5	16.6	109	2	C30608	Ig kappa chain V-I
37	441.5	16.5	129	2	S46369	Ig light chain V r
38	441.5	16.5	134	2	S38643	Ig kappa chain V r
39	440.5	16.5	107	2	PH0965	Ig kappa chain V r
40	440.5	16.5	109	2	PH0963	Ig kappa chain V r
41	438.5	16.4	109	2	A30608	Ig kappa chain V-I
42	438.5	16.4	109	2	F30607	Ig kappa chain V-I
43	438	16.4	107	2	S57444	Ig kappa chain V-J
44	437.5	16.4	109	2	B30601	Ig kappa chain V-I
45	436.5	16.4	109	2	G30601	Ig kappa chain V-I

#### ALIGNMENTS

RESULT 1  
RWHUT4  
T-cell surface glycoprotein CD4 precursor [validated] - human  
N.Alternate names: T-cell surface antigen T4/Leu 3  
C.Species: Homo sapiens (man)  
C.Date: 28-May-1986 #sequence\_revision 31-Dec-1988 #text\_change 09-Jul-2004  
C.Accession: A90872; A32722; A31194; A53287; I54176; I54297; A02109; A30039  
R.Maddon, P.J.; Littman, D.R.; Godfrey, M.; Maddon, D.E.; Chess, L.; Axel, R.  
Cell 42, 93-104, 1985  
A.Title: The isolation and nucleotide sequence of a cDNA encoding the T cell surface pro  
A.Reference number: A90872; MUID:85254948; PMID:2990730  
A.Accession: A90872  
A.Molecule type: mRNA  
A.Residues: 1-25, 'N', 27-458 <MAD>  
A.Cross-references: UNIPROT:P01730  
A.Experimental source: clone pT4B  
R.Littman, D.R.; Maddon, P.J.; Axel, R.  
Cell 55, 541, 1988  
A.Title: Corrected CD4 sequence.  
A.Reference number: A90907; MUID:89028665; PMID:3263213  
A.Contents: annotation; revision to residue 26  
R.Camerini, D.; Seed, B.  
Cell 60, 747-754, 1990  
A.Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi  
A.Reference number: A32722; MUID:90182664; PMID:2107024  
A.Accession: A32722  
A.Molecule type: mRNA  
A.Residues: 26-426, 428-458 <CAM>  
R.Carr, S.A.; Hemling, M.E.; Folea-Maserman, G.; Sweet, R.W.; Anumula, K.; Barr, J.R.;  
J. Biol. Chem. 264, 21286-21295, 1989  
A.Title: Protein and carbohydrate structural analysis of a recombinant soluble CD4 recep  
A.Reference number: A34194; MUID:90078332; PMID:2593374  
A.Contents: disulfide bonds; carbohydrate-binding sites  
A.Accession: A34194  
A.Molecule type: Protein  
A.Residues: 26-394 <CAR>  
R.Liederman, S.; DeMartino, J.A.; Daugherty, B.L.; Foeldvari, I.; Yellin, M.J.; Cleary, A  
Mol. Immunol. 28, 1171-1181, 1991  
A.Title: A single amino acid substitution in a common African allele of the CD4 molecule  
A.Reference number: A53287; MUID:92072595; PMID:1961196  
A.Status: not compared with conceptual translation  
A.Molecule type: mRNA  
A.Residues: 250-264, 'W', 266-280 <LED>  
A.Note: sequence extracted from NCBI backbone (NCBI:68249)  
R.Edwards, M.C.; Gibbs, R.A.  
Genome 14, 590-597, 1992  
A.Title: A human dimorphism resulting from loss of an Alu.  
A.Reference number: I54176; MUID:93052387; PMID:1330888  
A.Accession: I54176  
A.Status: translated from GR/EMBL/DBJ

A/Molecule type: DNA  
 A/Residues: 1-72 <RB5>  
 A/Cross-references: GB:U47924; GB:M86525; GB:U72506; NID:G1633547; PIDN:AA51309.1; PID:  
 R/hodge, T.W.; Sasso, D.R.; McDougal, J.S.  
 Hum. Immunol. 30, 99-104, 1991  
 A/Title: Humans with OKT4-epitope deficiency have a single nucleotide base change in the  
 A/Reference number: I54297; MUID:91216786; PMID:11708753  
 A/Accession: I54297  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-264, 'W', 266-458 <RB2>  
 A/Cross-references: GB:M35160; NID:G179143; PIDN:AA16069.1; PID:G179144  
 A/Comment: Macrophage tropic strains of HIV-1 bind to a complex of chemokine (C-C) recep  
 C/Genetics:  
 A/Gene: GDB:CD4  
 A/Cross-references: GDB:119767; OMIM:186940  
 A/Map position: 12pter-12p12  
 A/Intons: 16/3  
 C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C/Keywords: AIDS; duplication; glycoprotein; T-cell; transmembrane protein  
 F/1-25/Domain: signal sequence #status predicted <SIG>  
 F/26-458/Product: T-cell surface glycoprotein CD4 #status experimental <MAT>  
 F/34-111/Domain: immunoglobulin homology <IM1>  
 F/136-186/Domain: immunoglobulin homology #status atypical <IM2>  
 F/216-299/Domain: immunoglobulin homology <IM3>  
 F/321-372/Domain: immunoglobulin homology <IM4>  
 F/397-420/Domain: transmembrane #status predicted <TM>  
 F/421-458/Domain: intracellular #status predicted <INT>  
 F/41-109,155-184,328-370/Disulfide bonds: #status experimental  
 F/286,325/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 39.9%; Score 1064; DB 1; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 1e-48;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALLPAATGKRVYLGKGDYELTCTASQKSIQFHMKNNOIK 60  
 DB 1 MNRGVPRHLLVQLALLPAATGKRVYLGKGDYELTCTASQKSIQFHMKNNOIK 60  
 QY 61 ILNQGSEFLTKGSPKLNDRSDRSRLMDQGNFPIILNKLIEDSDTYICEVEDQKEEVOL 120  
 DB 61 ILNQGSEFLTKGSPKLNDRSDRSRLMDQGNFPIILNKLIEDSDTYICEVEDQKEEVOL 120  
 QY 121 LVEGLTANSDTHLLQGSLTTLTSPGSSPSVOCRSRPRGNIOGKTLVSQLELDSG 180  
 DB 121 LVEGLTANSDTHLLQGSLTTLTSPGSSPSVOCRSRPRGNIOGKTLVSQLELDSG 180  
 QY 181 TWCTVLQNKQVEFKIDIVLAFQKAS 208  
 DB 181 TWCTVLQNKQVEFKIDIVLAFQKAS 208

## RESULT 2

RWCZT4  
 T-cell surface glycoprotein CD4 - chimpanzee  
 N/Alternate names: T-cell surface antigen T4/Leu 3  
 C/Species: Pan troglodytes (chimpanzee)  
 C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C/Accession: B32722; A46534  
 R/Camerini, D.; Seed, B.  
 Cell 60, 747-754, 1990  
 A/Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi  
 A/Reference number: A32722; MUID:90182664; PMID:2107024  
 A/Accession: B32722  
 A/Molecule type: mRNA  
 A/Residues: 1-432 <CAM>  
 A/Cross-references: UNIPROT:P16004; GB:M31135  
 R/Pomsgaard, A.; Hirsch, V.M.; Johnson, P.R.  
 Eur. J. Immunol. 22, 2973-2981, 1992  
 A/Title: Cloning and sequences of primate CD4 molecules: diversity of the cellular recep  
 A/Reference number: A46534; MUID:93046640; PMID:11425921  
 A/Accession: A46534  
 A/Status: not compared with conceptual translation

A/Molecule type: mRNA  
 A/Residues: 3-399 <FOM>  
 A/Note: sequence extracted from NCBI backbone (NCBIP:118332)  
 C/Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells ct  
 C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C/Keywords: duplication; glycoprotein; T-cell; transmembrane protein  
 F/1-432/Product: T-cell surface glycoprotein CD4 #status predicted <MAT>  
 F/1-373/Domain: extracellular #status predicted <EXT>  
 F/9-86/Domain: immunoglobulin homology <IM1>  
 F/111-161/Domain: immunoglobulin homology #status atypical <IM2>  
 F/191-274/Domain: immunoglobulin homology <IM3>  
 F/286-347/Domain: immunoglobulin homology <IM4>  
 F/372-395/Domain: transmembrane #status predicted <TM>  
 F/396-432/Domain: intracellular #status predicted <INT>  
 F/16-84,130-159,303-345/Disulfide bonds: #status predicted  
 F/271,300/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 34.3%; Score 916; DB 1; Length 432;  
 Best Local Similarity 97.8%; Pred. No. 4.9e-41;  
 Matches 179; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 26 KKVVLGKGDYELTCTASQKSIQFHMKNNOIKILNQGSEFLTKGSPKLNDRSDRS 85  
 DB 1 KKVVLGKGDYELTCTASQKSIQFHMKNNOIKILNQGSEFLTKGSPKLNDRSDRS 85  
 QY 86 LMDQGNFPIILNKLIEDSDTYICEVEDQKEEVOLVFGTLTANSDTHLLQGSLTTLTLES 145  
 DB 86 LMDQGNFPIILNKLIEDSDTYICEVEDQKEEVOLVFGTLTANSDTHLLQGSLTTLTLES 120  
 QY 146 PPGSSPSVOCRSRPRGNIOGKTLVSQLELDSGWTCTVLQNKQVEFKIDIVLAFQ 205  
 DB 121 PPGSSPSVOCRSRPRGNIOGKTLVSQLELDSGWTCTVLQNKQVEFKIDIVLAFQ 180  
 QY 206 KAS 208  
 DB 181 KAS 183

## RESULT 3

RWMOZ4  
 T-cell surface glycoprotein CD4 - rhesus macaque  
 N/Alternate names: T-cell surface antigen T4/Leu 3  
 C/Species: Macaca mulatta (rhesus macaque)  
 C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Jul-1999  
 C/Accession: C32722  
 R/Camerini, D.; Seed, B.  
 Cell 60, 747-754, 1990  
 A/Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi  
 A/Reference number: A32722; MUID:90182664; PMID:2107024  
 A/Accession: C32722  
 A/Molecule type: mRNA  
 A/Residues: 1-432 <CAM>  
 A/Cross-references: GB:M31134  
 C/Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells ct  
 C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C/Keywords: duplication; glycoprotein; T-cell; transmembrane protein  
 F/1-432/Product: T-cell surface glycoprotein CD4 #status predicted <MAT>  
 F/1-373/Domain: extracellular #status predicted <EXT>  
 F/9-86/Domain: immunoglobulin homology <IM1>  
 F/111-161/Domain: immunoglobulin homology #status atypical <IM2>  
 F/180-293/Domain: immunoglobulin homology <IM3>  
 F/286-347/Domain: immunoglobulin homology <IM4>  
 F/372-395/Domain: transmembrane #status predicted <TM>  
 F/396-432/Domain: intracellular #status predicted <INT>  
 F/16-84,130-159,303-345/Disulfide bonds: #status predicted  
 F/271,300/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 31.1%; Score 831; DB 1; Length 432;  
 Best Local Similarity 88.0%; Pred. No. 1.3e-36;  
 Matches 161; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY 26 KKVVLGKGDYELTCTASQKSIQFHMKNNOIKILNQGSEFLTKGSPKLNDRSDRS 85  
 DB 1 KKVVLGKGDYELTCTASQKSIQFHMKNNOIKILNQGSEFLTKGSPKLNDRSDRS 85

Db 1 KKVVLGKGGDTVELTCTASQKNTQPFHMKNSNQIKILIGLFLTKGPKLSDRADSRKS 60  
 QY 86 LMDQNFPLIINKLKIEDSDTYICEVEDQKEVQLLVREGLTANSPDTHLLQGSLLTLES 145  
 Db 61 LMDQCFSPHITINKLKIEDSDTYICEVENKEVEVLLVPEGLTANSPDTHLLQGSLLTLES 120  
 QY 146 PGGSSPSVQCRRPRGNKIQGKTLVSQLELQDSGTWCTVYLNQKVEFKIDIVLAFO 205  
 Db 121 PGGSSPSVQCRSPGKNIQGGRTISVPQERODSGTWCTVSDQKTVFEKIDIVLAFO 180  
 QY 206 KAS 208  
 Db 181 KAS 183

## RESULT 4

A:Accession: A56446  
 Ig heavy chain V region (3H-3H scFv) - mouse (strain BALB/c)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 19-Jan-1996 #sequence\_revision 19-Jan-1996 #text\_change 16-Aug-1996  
 C:Accession: A56446  
 R:Yang, F.M.; Foltz, L.A.; Mahoney, W.C.; Schueler, P.A.  
 J. Biol. Chem. 270, 7829-7835, 1995  
 A:Title: A high affinity digoxin-binding protein displayed on M13 is functionally identical  
 A:Reference number: A56446; MUID:95229583; PMID:7713873  
 A:Accession: A56446  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-268 <TAN>  
 A:Cross-references: GB:U20617  
 C:Keywords: heterotrimer; immunoglobulin

Query Match 28.1%; Score 750; DB 2; Length 268;  
 Best Local Similarity 55.6%; Pred. No. 1.4e-32;  
 Matches 150; Conservative 39; Mismatches 57; Indels 24; Gaps 7;  
 QY 243 SOVQLSGAEVKKGGSSVSKVSCXASG---DTFIRYSTFWROAPGQLEMMGRITLL 298  
 Db 2 AOVKLQSGAEVKKGGSSVSKVSCXASG---DTFIRYSTFWROAPGQLEMMGRITLL 298  
 QY 239 DVAAHYAPHLQGRVTITADKSTSTVYLELNLSDDTAVYFCAGVYEGEADGEYDNNGL 358  
 Db 58 GITKIDPKFGKATIAADTSSNTAVYLGSLSTSEDTAVYCYASY---LTRYEN---- 108  
 QY 339 KMGCGTLVTYSSGGSGSGSGSGSGSELELTQSPATLSVSPGERPATLSGASASVSDD 418  
 Db 109 -YMGCGTLVTYSSGGSGSGSGSGSGSDIELTQSPATLSGSEKVTMSCRASSSVNF 166  
 QY 419 LAMVQKSGQAPRLIYGASTRATGVPARFSGSGSGABFTLTISSLSQSEDFAVYCCQOYN 478  
 Db 167 IYVQKSGDASPKLWYITTSHPGVPAPFSGSGSGSNSTLTISSMEGDAATYCCQOFT 226  
 QY 479 NMPRYTGGQTRLEIKLVPRSGGHHHHH 508  
 Db 227 SSP--FTFGSGTKLEIK---RSAHHHHH 251

## RESULT 5

S41374  
 single chain Fv antibody - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 06-Jan-1995  
 C:Accession: S41374  
 R:Artsaenko, O.; Weiler, E.W.; Wuentz, K.; Conrad, U.  
 submitted to the EMBL Data Library, January 1994  
 A:Description: Construction and functional characterization of a single chain Fv antibody  
 A:Reference number: S41374  
 A:Accession: S41374  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-249 <ART>  
 A:Cross-references: EMBL:Z29480

Query Match 25.0%; Score 668; DB 2; Length 249;  
 Best Local Similarity 52.9%; Pred. No. 2.5e-28;  
 Matches 136; Conservative 40; Mismatches 65; Indels 16; Gaps 4;

QY 244 QVQLLESGAEVKKGGSSVSKVSCXASGDTFIRYSTFWROAPGQLEMMGRITLLDVAAHY 303  
 Db 1 QVQLQSGAEVLRPAASVYKLSCTAGSNFQKODIYHWVQRPKEGLWIAIRIAPASGNVXY 60  
 QY 304 APHLQGRVTITADKSTSTVYLELNLSDDTAVYFCAGVYEGEADGEYDNNGLKHMGO 363  
 Db 61 VPRFQDKATITADTSSNTAVYLGSLSTSEDTAVYCYA-----RDITLTSTIGY--WGQ 111  
 QY 364 GTLVVTSGGGSGSGSGSGSELELTQSPATLSVSPGERPATLSGASASVS---SSD 418  
 Db 112 GSTVTSRRGGSGSGSGSGSDIELTQSPSVVIRFESVYSISCRSSKLLYSDDGSY 171  
 QY 419 LAMVQKSGQAPRLIYGASTRATGVPARFSGSGSGAEFTLLTSSLSQSEDFAVYCCQOYN 478  
 Db 172 LFMFLQRBQSGPQLLIYMSNLASGVPRFSGSGSGTFTLIRISRVAEADGVYCCMQHR 231  
 QY 479 NMPRYTGGQTRLEIK 495  
 Db 232 EYP--LTFGAGTKLEIK 246

## RESULT 6

A:Accession: A46254  
 CD4 precursor - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
 C:Accession: A46254  
 R:Hague, B.F.; Sawadkoseol, S.; Brown, T.U.; Lee, K.; Recker, D.P.; Kindt, T.J.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 7963-7967, 1992  
 A:Title: CD4 and its role in infection of rabbit cell lines by human immunodeficiency vi  
 A:Note: sequence extracted from NCBI backbone (NCBI:112732, NCBI:112733)  
 C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 F:322-372/Domain: immunoglobulin homology <IMM>

Query Match 23.2%; Score 618.5; DB 2; Length 459;  
 Best Local Similarity 59.6%; Pred. No. 1.7e-25;  
 Matches 127; Conservative 34; Mismatches 47; Indels 5; Gaps 2;

QY 1 MNRGVPFHLILVLDLALLPAATGKRVLGKGDVTELTCTASQKSIQFHMKNSNQIK 60  
 Db 1 MNRRIYFQCLLILVPLALLPAATGKTVVRGAGIYELPCQSSQKRSVFNMKHANVYK 60  
 QY 61 ILGNNG----SFLTKGPKLMDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKE 116  
 Db 61 ILGNQSSSSSFWLKGNSPLNRYESKKNMMDQSFPLVITLDLMDSDGTYICEVEDQK 120  
 QY 117 EVQLLVFGLTANSPDTHLLQGSLLTLESPPGSSPSVQCRRPRGNKIQGKTLVSQLEL 176  
 Db 121 EVELLVFRLTANPMTFRLHGGSLTLLEGPSVSVQWPKSKPKKIETGTCSPKRL 180  
 QY 177 QDSGTWCTV-LQNKQKVEFKIDIVLAFOKAS 208  
 Db 181 QDSGTWCTV-LQNKQKVEFKIDIVLAFOKAS 213

## RESULT 7

JC5322  
 p53 specific single-chain antibody Pab421 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 15-May-1997 #sequence\_revision 15-May-1997 #text\_change 18-Jul-1997  
 C:Accession: JC5322  
 R:Janot, C.B.; Hynes, N.E.  
 Biochem. Biophys. Res. Commun. 230, 242-246, 1997



Db

81 RFGSGSGTEFTLTISLSQSEDFAVYYCOQYNWNP--YTFGQGTKLEIK 127

## RESULT 11

I: kappa chain V region - human  
C: Species: Homo sapiens (man)  
C: Date: 02-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 21-Jan-2000  
C: Accession: S34005; S30524  
R: Mariette, X.; Tsapis, A.; Brouet, J.C.  
E: Eur. J. Immunol. 23, 846-851, 1993  
A: Title: Nucleotide sequence analysis of the variable domains of four human monoclonal  
A: Reference number: S34001; MUID:9330281; PMID:768198  
A: Accession: S34005  
A: Status: preliminary  
A: Molecule type: mRNA  
A: Residues: 1-107 <MAR>  
A: Cross-references: EMBL: Z18330  
C: Superfamily: immunoglobulin V region; immunoglobulin homology  
C: Keywords: heterotetramer; immunoglobulin  
E: 16-90/Domain: immunoglobulin homology <IWM>

## RESULT 12

Ig kappa chain - human  
C/Species: Homo sapiens (man)  
C/Date: 06-Mar-1994 #sequence\_revision 26-May-1995 #text\_change 21-Jan-2000  
C/Accession: SA0362  
R/Klein, R.; Jaenichen, R.; Zachau, H.G.  
Eur. J. Immunol. 23, 3248-3271, 1993  
A/Title: Expressed human immunoglobulin chi genes and their hypermutation.  
A/Reference number: S40312; M0ID:94080891; PMID:8258341  
A/Accession: SA0362  
A/Status: preliminary; translation not shown  
A/Molecule type: mRNA  
A/Residues: 1-117 <KLE>  
A/Cross-references: EMBL:X72472; NID:G441412; PID:G441413  
C/Superfamily: immunoglobulin V region; immunoglobulin homology  
C/Keywords: heterotetramer; immunoglobulin  
E/26-100/Domain: immunoglobulin homology <IMM>

## RESULT 13

T-cell surface glycoprotein CD4 precursor - mouse  
N;Alternate names: T-cell differentiation antigen L3T4; T-cell surface antigen T4/Leu 3  
N;Species: Mus musculus (house mouse)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:28:31 ; Search time 281.81 Seconds  
(without alignments)  
1037.189 Million cell updates/sec

Title: US-09-936-702-3

Perfect score: 2668

Sequence: 1 MNRGVPRHLLVLQLALLP.....GTRLEIKLVPRGSHHHHH 508

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 02: \*  
1: uniprot\_sprot: \*  
2: uniprot\_crembl: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1064	39.9	458	1	CD4_HUMAN
2	1040	39.0	458	1	CD4_PANTR
3	953	35.7	458	1	CD4_MACFA
4	951	35.6	458	1	CD4_MACFU
5	945	35.4	458	1	CD4_MACMU
6	944	35.4	458	1	CD4_MACNE
7	926	34.7	458	1	CD4_CERAB
8	820	30.7	397	1	CD4_CERTO
9	819	30.7	397	1	CD4_ERKPA
10	803	30.1	397	2	009261
11	803	30.1	397	2	009262
12	803	30.1	397	2	009263
13	803	30.1	397	2	009264
14	799	29.9	397	2	009259
15	799	29.9	397	2	009260
16	778.5	29.2	457	2	08H2T8
17	777.5	29.1	457	2	08H2T7
18	774.5	29.0	457	1	CD4_SAIISC
19	766	28.7	255	2	06K805
20	766	28.7	255	2	CAG34081
21	750	28.1	243	2	Q7T0M2
22	734	27.5	298	2	Q9GYF0
23	685.5	25.7	241	2	0921A6
24	632	23.7	463	1	CD4_CANFA
25	630	23.6	455	2	0710E2
26	630	23.6	455	2	AAQ03208
27	625	23.4	455	2	Q9XS78
28	618.5	23.2	459	1	CD4_RABIT
29	612	22.3	218	2	Q925S1
30	596	22.3	432	2	06L8K1
31	596	22.3	432	2	CAB37664

32	590.5	22.1	474	2	P79355	P79355 felis silve
33	585	21.9	417	2	AA99809	AA99809 sus scrofa
34	585	21.9	457	2	Q6GYR3	Q6GYR3 sus scrofa
35	585	21.9	457	2	AA99808	AA99808 sus scrofa
36	563	21.1	406	2	Q6R3N3	Q6R3N3 sus scrofa
37	563	21.1	447	2	Q6R3N4	Q6R3N4 sus scrofa
38	508.5	19.1	457	1	CD4_RAT	P05540 rattus norv
39	505.5	18.9	129	1	KV3H_HUMAN	P04207 homo sapien
40	498	18.7	457	1	CD4_MOUSE	P06332 mus musculu
41	490	18.4	457	2	Q61396	Q61396 mus musculu
42	485.5	18.2	109	2	Q9UL85	Q9UL85 homo sapien
43	480.5	18.0	109	1	KV3F_HUMAN	P01624 homo sapien
44	480	18.0	108	2	Q9UL83	Q9UL83 homo sapien
45	480	18.0	234	2	AAH30813	AAH30813 homo sapi

## ALIGNMENTS

RESULT 1  
CD4\_HUMAN STANDARD; PRT; 458 AA.  
AC P01730;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 01-OCT-2004 (Rel. 45, Last annotation update)  
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen  
DE T4/Leu-3).  
GN Name=CD4;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OK NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85254948; PubMed=2990730;  
RA Maddon P.J., Littman D.R., Godfrey M., Maddon D.B., Chess L., Axel R.,  
RT "The isolation and nucleotide sequence of a cDNA encoding the T cell,  
RT surface protein T4: a new member of the immunoglobulin gene family.";  
RL Cell 42:93-104 (1985).  
RN [2]  
RP REVISION TO 26.  
RX MEDLINE=89028655; PubMed=3263213;  
RA Littman D.R., Maddon P.J., Axel R.,  
RT "Corrected CD4 sequence.";  
RL Cell 55:541-541 (1988).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96303695; PubMed=8723724;  
RA Anselmi-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E., Spanos S.,  
RT "A gene-rich cluster between the CD4 and triosephosphate isomerase  
RT genes at human chromosome 12p13.";  
RL Genome Res. 6:314-326 (1996).  
RN [4]  
RP SEQUENCE FROM N.A., AND VARIANT TRP-265.  
RX MEDLINE=91216786; PubMed=1708753;  
RA Hodge T.W., Sasse D.R., McDougal J.S.,  
RT "Humans with OKT4-epitope deficiency have a single nucleotide base  
RT change in the CD4 gene, resulting in substitution of TRP240 for  
RT ARG240.";  
RL Hum. Immunol. 30:99-104 (1991).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Pancreas;  
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Martina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Moxley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Finley J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schenck A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [6]  
 RP SEQUENCE OF 28-424 FROM N.A.  
 RC TISSUE=Blood;  
 RX MEDLINE=93049640; PubMed=1425921.  
 RA Pomegaard A., Hirsch V.M., Johnson P.R.,  
 RT "Cloning and sequences of primate CD4 molecules: diversity of the  
 RT cellular receptor for simian immunodeficiency virus/human  
 RT immunodeficiency virus.";  
 RL Eur. J. Immunol. 22:2973-2981(1992).  
 RN [7]  
 RP SEQUENCE OF 26-394.  
 RX MEDLINE=90078232; PubMed=2592374;  
 RA Carr S.A., Hemling M.E., Folea-Wasserman G., Sweet R.W., Anumula K.,  
 RA Barr J.R., Huddleston M.J., Taylor P.,  
 RT "Protein and carbohydrate structural analysis of a recombinant soluble  
 RT CD4 receptor by mass spectrometry.";  
 RL J. Biol. Chem. 264:21286-21295(1989).  
 RN [8]  
 RP SEQUENCE OF 26-40.  
 RA Zhang Z., Henzel W.,  
 RT "Signal peptide prediction based on analysis of experimentally  
 RT verified cleavage sites.";  
 RL Submitted (JUN-2004) to Swiss-Prot.  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 26-208.  
 RX MEDLINE=91061881; PubMed=1701030;  
 RA Wang J., Yan Y., Garrett T.P., Liu J., Rodgers D.W., Garlick R.L.,  
 RA Tarr G.E., Hsuan Y., Reinherz E.L., Harrison S.C.,  
 RT "Atomic structure of a fragment of human CD4 containing two  
 RT immunoglobulin-like domains.";  
 RL Nature 348:411-418(1990).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 26-208.  
 RX MEDLINE=91061882; PubMed=2247146;  
 RA Ryu S.-E., Kwong P.D., Truneh A., Porter T.G., Arthos J.,  
 RA Rosenberg M., Dai X., Xuong N.-H., Axel R., Sweet R.W.,  
 RA Hendrickson W.A.,  
 RT "Crystal structure of an HIV-binding recombinant fragment of human  
 RT CD4.";  
 RL Nature 348:419-426(1990).  
 RN [11]  
 RP X-RAY CRYSTALLOGRAPHY (3.9 ANGSTROMS) OF 26-388.  
 RX MEDLINE=97311402; PubMed=9168119;  
 RA Wu H., Kwong P.D., Hendrickson W.A.,  
 RT "Dimeric association and segmental variability in the structure of  
 RT human CD4.";  
 RL Nature 387:527-530(1997).  
 RN [12]  
 RP PALMITOYLATION.  
 RX MEDLINE=92317088; PubMed=1618861;  
 RA Crise B., Rose J.K.,  
 RT "Identification of palmitoylation sites on CD4, the human  
 RT immunodeficiency virus receptor.";  
 RL J. Biol. Chem. 267:13593-13597(1992).  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD4 entry;  
 CC WWW=http://www.ncbi.nlm.nih.gov/prow/cd/cd4.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; M12807; AAA35572.1; -;  
 DR EMBL; U47924; AAB51309.1; -;  
 DR EMBL; M35160; AAA16069.1; -;  
 DR EMBL; BC025782; AAH25782.1; -;  
 DR PIR; A90872; RWHUT4.  
 DR PDB; 1CDH; X-ray; @=26-203.  
 DR PDB; 1CDI; X-ray; @=25-203.  
 DR PDB; 1CDJ; X-ray; @=26-203.  
 DR PDB; 1CDU; X-ray; @=26-203.  
 DR PDB; 1CDY; X-ray; @=26-203.  
 DR PDB; 1G9M; X-ray; @=26-210.  
 DR PDB; 1G9N; X-ray; C=26-210.  
 DR PDB; 1GCI; X-ray; C=26-210.  
 DR PDB; 1JL4; X-ray; D=26-203.  
 DR PDB; 1Q68; NMR; A=421-458.  
 DR PDB; 1WBR; NMR; @=427-445.  
 DR PDB; 1WIP; X-ray; A/B=26-388.  
 DR PDB; 1WIO; X-ray; A/B=26-388.  
 DR PDB; 3CD4; X-ray; @=26-207.  
 DR IntAct; P01730; -;  
 DR GlycoSuiteDB; P01730; -;  
 DR Genew; HGNC:1678; CD4.  
 DR MIM; 186340; -;  
 DR GO; GO:0005886; C:Plasma membrane; TAS.  
 DR GO; GO:0042101; C:T-cell receptor complex; NAS.  
 DR GO; GO:0015026; F:coreceptor activity; NAS.  
 DR GO; GO:0015029; F:internalization receptor activity; TAS.  
 DR GO; GO:0042389; F:MHC class II protein binding; NAS.  
 DR GO; GO:0004888; F:transmembrane receptor activity; TAS.  
 DR GO; GO:0006955; P:immune response; NAS.  
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; NAS.  
 DR GO; GO:0030217; P:T-cell differentiation; NAS.  
 DR GO; GO:0045058; P:T-cell selection; NAS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; NAS.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TCRg.  
 DR InterPro; IPR007110; Ig\_1like.  
 DR InterPro; IPR003596; Ig\_V.  
 DR Pfam; PF00047; Ig\_2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV\_1.  
 DR PROSITE; PS50835; IG\_1like; 1.  
 KW 3D-structure; Direct protein sequencing; Glycoprotein;  
 KW Immune response; Immunoglobulin domain; Lipoprotein; Palmitate;  
 KW Polymorphism; Repeat; Signal; T-cell; Transmembrane.  
 FT SIGNAL 1 25  
 FT CHAIN 26 458  
 FT DOMAIN 26 396  
 FT TRANSMEM 397 418  
 FT DOMAIN 419 458  
 FT DOMAIN 26 125  
 FT DOMAIN 126 203  
 FT DOMAIN 204 317  
 FT DOMAIN 318 374  
 FT CARBOHYD 296  
 FT CARBOHYD 325  
 FT DISULFID 41 109  
 FT DISULFID 155 184  
 FT DISULFID 328 370  
 By similarity.



FT LIPID 419 419 S-palmitoyl cysteine.  
 FT LIPID 422 422 S-palmitoyl cysteine.  
 Query Match 39.9%; Score 1064; DB 1; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-59;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRLHLVLTALLPATQGGKVVGLGKKGDVTELCTASQKKSIOFHWKNSNOTK 60  
 DB 1 MNRGVPFRLHLVLTALLPATQGGKVVGLGKKGDVTELCTASQKKSIOFHWKNSNOTK 60  
 QY 61 ILNGGSPFLTKGSPSKLNDRADSRSLMDQGNFPLIKKLIKEDSDTYICEVEDQKEEYOL 120  
 DB 61 ILNGGSPFLTKGSPSKLNDRADSRSLMDQGNFPLIKKLIKEDSDTYICEVEDQKEEYOL 120  
 QY 121 LVFGLTANSDFHLQGGSLTLTLSPSPSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDFHLQGGSLTLTLSPSPSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWTCTVLONQKVKERKIDIVLAFOKAS 208  
 DB 181 TWTCTVLONQKVKERKIDIVLAFOKAS 208

RESULT 2  
 ID CD4\_PANTR STANDARD; PRT; 458 AA.  
 AC P16004;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).  
 GN Name=CD4;  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
 OX NCBI\_TaxID=9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=90182664; PubMed=2107024;  
 RA Camerini D., Seed B.;  
 RT "A CD4 domain important for HIV-mediated syncytium formation lies outside the virus binding site.";  
 RL Cell 60:747-754 (1990).  
 RN [2]  
 RP SEQUENCE OF 26-424 FROM N.A.  
 RC TISSUE=Blood;  
 RX MEDLINE=93049640; PubMed=1425921;  
 RA Pomegaard A., Hirsch V.M., Johnson P.R.;  
 RT "Cloning and sequences of primate CD4 molecules: diversity of the cellular receptor for simian immunodeficiency virus/human immunodeficiency virus.";  
 RL Eur. J. Immunol. 22:2973-2981 (1992).  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL, M31135; AA35407.1; -  
 DR EMBL, X73353; CA551749.1; -  
 DR PIR, B32722; RWC274.  
 DR HSPB, P01730; 1W10.

DR GO:0042101; C-T-cell receptor complex; ISS.  
 DR GO:0015026; Fc-gamma receptor activity; ISS.  
 DR GO:0042289; Fc-gamma class II protein binding; ISS.  
 DR GO:0006955; P:immune response; ISS.  
 DR GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.  
 DR GO:0030217; P:T-cell differentiation; ISS.  
 DR GO:0045058; P:T-cell selection; ISS.  
 DR GO:0007168; P:transmembrane receptor protein tyrosine kin. . .; ISS.  
 DR InterPro: IPR008424; CD2.  
 DR InterPro: IPR009713; CD4\_TcAg.  
 DR InterPro: IPR007110; Ig\_Like.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam: PF00047; Ig\_2.  
 DR PRINTS: PR00692; CD4TCANTIGEN.  
 DR SMART: SM00406; IGV\_1.  
 DR PROSITE: PS00835; IG\_LIKE; 1.  
 DR KX Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;  
 KW Palmitate; Repeat; Signal; T-cell; Transmembrane.  
 FT SIGNAL 1 25  
 FT CHAIN 26 458  
 FT DOMAIN 26 396  
 FT TRANSMEM 397 418  
 FT DOMAIN 419 458  
 FT DOMAIN 126 203  
 FT DOMAIN 204 317  
 FT DOMAIN 318 374  
 FT CARBOHYD 296 296  
 FT CARBOHYD 325 325  
 FT DISULFID 41 109  
 FT DISULFID 155 184  
 FT DISULFID 328 370  
 FT LIPID 419 419  
 FT LIPID 422 422  
 FT CONFLICT 42 42  
 FT CONFLICT 62 62  
 FT CONFLICT 191 191  
 SQ SEQUENCE 458 AA; 51057 MW; A7C3AC8A5257D3AD CRC64;

Query Match 39.0%; Score 1040; DB 1; Length 458;  
 Best Local Similarity 98.1%; Pred. No. 2e-57;  
 Matches 204; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MNRGVPFRLHLVLTALLPATQGGKVVGLGKKGDVTELCTASQKKSIOFHWKNSNOTK 60  
 DB 1 MNRGVPFRLHLVLTALLPATQGGKVVGLGKKGDVTELCTASQKKSIOFHWKNSNOTK 60  
 QY 61 ILNGGSPFLTKGSPSKLNDRADSRSLMDQGNFPLIKKLIKEDSDTYICEVEDQKEEYOL 120  
 DB 61 ILNGGSPFLTKGSPSKLNDRADSRSLMDQGNFPLIKKLIKEDSDTYICEVEDQKEEYOL 120  
 QY 121 LVFGLTANSDFHLQGGSLTLTLSPSPSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDFHLQGGSLTLTLSPSPSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWTCTVLONQKVKERKIDIVLAFOKAS 208  
 DB 181 TWTCTVLONQKVKERKIDIVLAFOKAS 208

RESULT 3  
 ID CD4\_MACFA STANDARD; PRT; 458 AA.  
 AC P79185;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).  
 GN Name=CD4;  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecinae; Macaca.  
 ON NCBI\_Taxid=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Thymocytes;  
 RA Tatusumi M., Yabe M., Yamada Y.K.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1 FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1 SUBUNIT: Associates with p56-lck (By similarity).  
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1 SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D63349; BAA09673.1; -.  
 DR HSSP; P01730; IMBR.  
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.  
 DR GO; GO:0006955; P:immune response; ISS.  
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.  
 DR GO; GO:0030217; P:T-cell differentiation; ISS.  
 DR GO; GO:0045058; P:T-cell selection; ISS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TcAg.  
 DR InterPro; IPR007110; Ig-Like.  
 DR InterPro; IPR003596; Ig\_V.  
 DR Pfam; PF00047; Ig\_2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;  
 KW Palmitate; Repeat; Signal; T-cell; Transmembrane.  
 FT SIGNAL 1 25  
 FT CHAIN 26 458  
 FT DOMAIN 26 396  
 FT TRANSEM 397 418  
 FT DOMAIN 419 458  
 FT DOMAIN 26 125  
 FT DOMAIN 126 203  
 FT DOMAIN 204 317  
 FT DOMAIN 318 374  
 FT CARBOHYD 42 42  
 FT CARBOHYD 296 296  
 FT CARBOHYD 325 325  
 FT DISULFID 41 109  
 FT DISULFID 155 184  
 FT DISULFID 328 370  
 FT LIPID 419 419  
 FT LIPID 422 422  
 FT SEQUENCE 458 AA; 50872 MW; 9105479FB5C56FF7 CRC64;  
 SQ  
 Query March 35.7%; Score 953; DB 1; Length 458;  
 Best Local Similarity 88.9%; Pred. No. 6e-52;  
 Matches 185; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 121 LVFGLTANSDPTLHLLQGSLTLLTLESPGSSPSVQCSPRGKNIQGGKTLTSSVSLQLQDSG 180  
 DB 121 LVFGLTANSDPTLHLLQGSLTLLTLESPGSSPSVQCSPRGKNIQGGKTLTSSVSLQLQDSG 180  
 QY 181 TWTCYLVQNKQKYEKFKIDIVLAFOKAS 208  
 DB 181 TWTCYLVQNKQKYEKFKIDIVLAFOKAS 208  
 RESULT 4  
 CD4\_MACFU STANDARD; PRT; 458 AA.  
 ID CD4\_MACFU  
 AC P79184;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 05-JUN-2004 (Rel. 44, Last annotation update)  
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen  
 T4/Leu-3).  
 GN Name=CD4;  
 OS Macaca fuscata fuscata (Japanese macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; Macaca.  
 ON NCBI\_Taxid=9543;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hashimoto O., Tatusumi M.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1 FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1 SUBUNIT: Associates with p56-lck (By similarity).  
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1 SIMILARITY: Contains 1 immunoglobulin-like C2-type domains.  
 CC -1 SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D63348; BAA09672.1; -.  
 DR HSSP; P01730; IMBR.  
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.  
 DR GO; GO:0006955; P:immune response; ISS.  
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.  
 DR GO; GO:0030217; P:T-cell differentiation; ISS.  
 DR GO; GO:0045058; P:T-cell selection; ISS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TcAg.  
 DR InterPro; IPR007110; Ig-Like.  
 DR InterPro; IPR003596; Ig\_V.  
 DR Pfam; PF00047; Ig\_2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;  
 KW Palmitate; Repeat; Signal; T-cell; Transmembrane.  
 FT SIGNAL 1 25  
 FT CHAIN 26 458  
 FT DOMAIN 26 396  
 FT TRANSEM 397 418  
 FT DOMAIN 419 458  
 FT DOMAIN 26 125  
 FT DOMAIN 126 203  
 FT DOMAIN 204 317  
 FT DOMAIN 318 374  
 FT CARBOHYD 42 42  
 FT CARBOHYD 296 296  
 FT CARBOHYD 325 325  
 FT DISULFID 41 109  
 FT DISULFID 155 184  
 FT DISULFID 328 370  
 FT LIPID 419 419  
 FT LIPID 422 422  
 FT SEQUENCE 458 AA; 50872 MW; 9105479FB5C56FF7 CRC64;  
 SQ  
 Query March 35.7%; Score 953; DB 1; Length 458;  
 Best Local Similarity 88.9%; Pred. No. 6e-52;  
 Matches 185; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

```

FT CARBOHYD 296 296 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 335 335 N-linked (GlcNAc...) (By similarity).
FT DISULFID 41 109 By similarity.
FT DISULFID 155 184 By similarity.
FT DISULFID 370 370 By similarity.
FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
SQ SEQUENCE 458 AA; 50828 MW; 76B3E7EF08185535 CRC64;

Query Match 35.6%; Score 951; DB 1; Length 458;
Best Local Similarity 88.0%; Pred. No. 8e-52;
Matches 184; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 NMRGVPFRHLVLTVALPALTATOGKKVYLKGGDVTVELTCTASOKKSIOFHMKNNOIK 60
DB 1 NMRGVPFRHLVLTVALPALTATOGKKVYLKGGDVTVELTCTASOKKSIOFHMKNNOIK 60
QY 61 ILGNQSPFLTKGPSKLNDRAISRSLMDQGNFPLIKNIKIEDSDTYICEVEDEKEEVL 120
DB 61 ILGIGSPFLTKGPSKLNDRAISRSLMDQGNFPLIKNIKIEDSDTYICEVEDEKEEVL 120
QY 121 LVFGITANSDFHLKQOSLTLTLESFPGSSPSVQCRSPGKNIQGGKTVLSQLELDQSG 180
DB 121 LVFGITANSDFHLKQOSLTLTLESFPGSSPSVQCRSPGKNIQGGKTVLSQLELDQSG 180
QY 181 TWTCTVLONOKVEPKIDIVLAFQKAS 208
DB 181 TWTCTVLONOKVEPKIDIVLAFQKAS 208
DB 181 TWTCTVLONOKVEPKIDIVLAFQKAS 208

RESULT 5
CD4_MACMU STANDARD; PRT; 458. AA.
ID CD4_MACMU STANDARD; PRT; 458. AA.
AC P16003; Q29617;
DT 01-APR-1990 (Rel. 14, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN Name=CD4;
OS Macaca mulatta (rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OC NCBI_TaxId=9544;
OX NCBI_TaxId=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90182664; PubMed=2107024;
RA Camerini D.; Seed B.;
RT "A CD4 domain important for HIV-mediated syncytium formation lies
RT outside the virus binding site.";
RL Cell 60:747-754(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Thymocytes;
RA Hashimoto O.; Tatsumi M.;
RT "Molecular cloning and expression of macaque CD4s.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 28-424 FROM N.A.
RX TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A.; Hirsch V.M.; Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus.";
RL Bur. J. Immunol. 22:2973-2981(1992).
RN [4]
RP SEQUENCE OF 107-192 FROM N.A.
RX MEDLINE=98320644; PubMed=9656488;
RA Harris E.E.; DiCoccia T.R.;
RT "Nuclear gene trees and the phylogenetic relationships of the
RT mangabeys (Primates: Papionini).";

```

```

RL Mol. Biol. Evol. 15:892-900(1998).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M31134; AAA6838.1; -.
CC EMBL; D63347; BAA09671.1; -.
CC EMBL; X73326; CAA51752.1; -.
CC EMBL; AF057385; AAC25129.1; -.
CC HSP; P01730; IWER.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn.; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin.; ISS.
DR InterPro; IPR008424; CD2.
DR InterPro; IPR009973; CD4_TGAG.
DR InterPro; IPR007110; Ig_V-like.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV_1.
DR PROSITE; PS50835; IGV_LIKE; 1.
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 25
FT CHAIN 26 458 T-cell surface glycoprotein CD4.
FT DOMAIN 26 396 Extracellular (Potential).
FT TRANSMEM 397 418 Potential.
FT DOMAIN 419 458 Cytoplasmic (Potential).
FT DOMAIN 126 125 Ig-1-like V-type.
FT DOMAIN 204 203 Ig-1-like C2-type 1.
FT DOMAIN 318 374 Ig-1-like C2-type 2.
FT DOMAIN 318 374 Ig-1-like C2-type 3.
FT CARBOHYD 296 296 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 325 325 N-linked (GlcNAc...) (By similarity).
FT DISULFID 41 109 By similarity.
FT DISULFID 155 184 By similarity.
FT DISULFID 370 370 By similarity.
FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
FT CONFLICT 42 42 N->T (in Ref. 1).
FT CONFLICT 62 62 L->S (in Ref. 3).
FT CONFLICT 67 67 L->S (in Ref. 2).
FT CONFLICT 169 169 I->L (in Ref. 2).
FT CONFLICT 191 191 K->N (in Ref. 2).
FT CONFLICT 248 248 S->P (in Ref. 2).
FT CONFLICT 265 265 R->Q (in Ref. 3).
FT CONFLICT 349 349 A->T (in Ref. 2).
SQ SEQUENCE 458 AA; 50884 MW; 8BB8039FAFEC808 CRC64;

Query Match 35.4%; Score 945; DB 1; Length 458;
Best Local Similarity 88.0%; Pred. No. 1.9e-51;
Matches 183; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 1 NMRGVPFRHLVLTVALPALTATOGKKVYLKGGDVTVELTCTASOKKSIOFHMKNNOIK 60
DB 1 NMRGVPFRHLVLTVALPALTATOGKKVYLKGGDVTVELTCTASOKKSIOFHMKNNOIK 60

```

```

QY 61 ILNGSGFLTKGSKLNDRASSRLMDQGNFPLIINKLIEPSDTYICEVEDQKEEVQL 120
DB 61 ILGIQGFLLTKRGSKLSDRADSRKSLMDQGFSLIINKLIEPSDTYICEVENKEEVEL 120
QY 121 LVFGLTANSTHLLQGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDG 180
DB 121 LVFGLTANSTHLLQGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDG 180
QY 181 TWICTVQNKQKVEFKIDIVLAFQKAS 208
DB 181 TWICTVQNKQKVEFKIDIVLAFQKAS 208

RESULT 6
CD4_MACNE STANDARD; PRT; 458 AA.
AC 008340; F79196;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN Name=CD4;
OS Macaca nemestrina (pig-tailed macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9545;
RN [1]
RP SEQUENCE FROM N.A.
RA Hashimoto O., Tatsumi M.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; Pubmed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus";
RT Eur. J. Immunol. 22:2973-2981(1992).
RL
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; D63346; BAA09670.1; -
DR EMBL; X73325; CA51751.1; -
DR HSSP; P01730; IMBR.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR000973; CD4_TcAg.
DR InterPro: IPR007110; Ig-Like.
DR InterPro: IPR003596; Ig_v.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.

```

```

DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG_Like; 1.
KM Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KM Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSSEM 397 418
FT DOMAIN 419 458
FT DOMAIN 426 458
FT DOMAIN 126 203
FT DOMAIN 317 374
FT DOMAIN 318 374
FT CARBOHYD 42 42
FT CARBOHYD 296 296
FT CARBOHYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT DISULFID 419 419
FT LIPID 422 422
FT CONFLICT 57 57
FT CONFLICT 91 91
FT CONFLICT 105 105
FT CONFLICT 113 113
FT CONFLICT 302 302
FT CONFLICT 349 349
SQ SEQUENCE 458 AA; 50905 MW; 751A9BA2CB83EBE6 CRC64;

Query Match 35.4%; Score 944; DB 1; Length 458;
Best Local Similarity 88.0%; Pred. No. 2, 2e-51;
Matches 183; Conservative 14; Mismatches 11; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLTLVQLALPAATGKKVYGGKGDVTELTCTSOXKSIQFHHKNSQIK 60
DB 1 MNRGVPFRHLTLVQLALPAATGKKVYGGKGDVTELTCTSOXKSIQFHHKNSQIK 60
QY 61 ILNGSGFLTKGSKLNDRASSRLMDQGNFPLIINKLIEPSDTYICEVEDQKEEVQL 120
DB 61 ILGIQGFLLTKRGSKLSDRADSRKSLMDQGFSLIINKLIEPSDTYICEVENKEEVEL 120
QY 121 LVFGLTANSTHLLQGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDG 180
DB 121 LVFGLTANSTHLLQGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDG 180
QY 181 TWICTVQNKQKVEFKIDIVLAFQKAS 208
DB 181 TWICTVQNKQKVEFKIDIVLAFQKAS 208

RESULT 7
CD4_CERAE STANDARD; PRT; 458 AA.
AC 008338; O02805; O77593; Q28217;
DT 01-FEB-1995 (Rel. 31, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN Name=CD4;
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RA Hashimoto O., Tatsumi M.;
RT "Molecular cloning and expression of african green monkey CD4.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Blood;

```

RX MEDLINE=93049640; PubMed=1425921;  
 RA Fomsgaard A., Hirsch V.M., Johnson P.R.;  
 RA "Cloning and sequences of primate CD4 molecules: diversity of the  
 RT cellular receptor for simian immunodeficiency virus/human  
 RT immunodeficiency virus.";  
 RL Eur. J. Immunol. 22:2973-2981(1992).  
 RL (3)  
 RP SEQUENCE OF 28-424 FROM N.A.  
 RC TISSUE=Peripheral blood;  
 RA MEDLINE=98017879; PubMed=9379478;  
 RA Fomsgaard A., Mueller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,  
 RA Corbet S., Barre-Sinoussi F., Allan J.S.;  
 RT "Relation between phylogeny of African green monkey CD4 genes and  
 RT their respective simian immunodeficiency virus genes.";  
 RL J. Med. Primatol. 26:120-128(1997).  
 RN (4)  
 RP SEQUENCE OF 107-192 FROM N.A.  
 RX MEDLINE=98320644; PubMed=9656488;  
 RA Harris E.E., Disotell T.R.;  
 RT "Nuclear gene trees and the phylogenetic relationships of the  
 RT mangabeys (Primates: Papionini).";  
 RL Mol. Biol. Evol. 15:892-900(1998).  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D86589; BAA13132.1; -  
 DR EMBL; X73322; CA51748.1; -  
 DR EMBL; AF001226; AAB60873.1; -  
 DR EMBL; AF001228; AAB60875.1; -  
 DR EMBL; AF057380; AAC25124.1; -  
 DR HSSP; P01730; IMIQ.  
 DR GO; GO:0042101; C-T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.  
 DR GO; GO:0006955; P:immune response; ISS.  
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.  
 DR GO; GO:0030217; P:T-cell differentiation; ISS.  
 DR GO; GO:0045058; P:T-cell selection; ISS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4 TCag.  
 DR InterPro; IPR007110; Ig-like.  
 DR Pfam; PF00047; Ig\_2.  
 DR PRINTS; PR00692; CD4TCANTIGN.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;  
 KW Palmitate; Repeat; Signal; T-cell; Transmembrane.  
 FT SIGNAL 1 25  
 FT CHAIN 26 458  
 FT DOMAIN 26 396  
 FT TRANSMEM 397 418  
 FT DOMAIN 419 458  
 FT DOMAIN 26 125  
 FT DOMAIN 126 203  
 FT DOMAIN 204 317  
 FT DOMAIN 318 374  
 FT CAROHD 42 42  
 FT CAROHD 281 281  
 FT CAROHD 296 296  
 FT CAROHD 325 325  
 FT DISULFID 41 109  
 By similarity.

FT DISULFID 155 184 By similarity.  
 FT DISULFID 328 370 By similarity.  
 FT LIPID 419 419 S-palmitoyl cysteine (By similarity).  
 FT LIPID 422 422 S-palmitoyl cysteine (By similarity).  
 FT CONFLICT 46 46 K -> N (in Ref. 2 and 3; AAB60875).  
 FT CONFLICT 59 59 I -> T (in Ref. 3; AAB60873).  
 FT CONFLICT 115 115 K -> E (in Ref. 1).  
 FT CONFLICT 165 165 G -> V (in Ref. 3; AAB60873 and 4).  
 FT CONFLICT 200 200 M -> V (in Ref. 2 and 3).  
 FT CONFLICT 227 227 F -> L (in Ref. 3; AAB60873).  
 FT CONFLICT 271 271 K -> E (in Ref. 3; AAB60873).  
 FT CONFLICT 281 281 N -> H (in Ref. 3; AAB60873).  
 SQ SEQUENCE 458 AA; 51158 MW; FC523D2EDD1F72E7 CRC64;  
 Query Match 34.7%; Score 926; DB 1; Length 458;  
 Best Local Similarity 86.5%; Pred. No. 3e-50;  
 Matches 180; Conservative 12; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 MNRGVPFRLILVLTALLPATQKRVLLGKGGTVELTCAQKKSIOFWKXNOIK 60  
 DB 1 MNWGIPEFRLILVLTALLPATQKRVLLGKGGTVELTCAQKKSIOFWKXNOIK 60  
 QY 61 ILKNGSFLTKGPSKLNDRADSRSLMPQGNPPLIILKULKIEDSPYICVEDOKEEVOL 120  
 DB 61 ILKNGSFLTKGSSLRDRIDSRKSLMPQGNPPLIILKULKIEDSETYICVENKKEEVL 120  
 QY 121 LVFGLTANSDFHLLQGQSLITLTLESPPGSSPEVQCRSPRKNIOGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDFHLLQGQSLITLTLESPPGSSPEVQCRSPRKNIOGKTLVSQLELDQSG 180  
 QY 181 TWTCTVLOKQKVEFKIDIVLAPQKAS 208  
 DB 181 TWTCTVLOKQKVEFKIDIVLAPQKAS 208  
 QY 181 TWTCTVLOKQKVEFKIDIVLAPQKAS 208  
 DB 181 TWTCTVLOKQKVEFKIDIVLAPQKAS 208  
 RESULT 8  
 ID CD4\_CERTO STANDARD; PRT; 397 AA.  
 AC 008336;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Leu-3)  
 GN (Fragment).  
 GN Name=CD4;  
 OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Butheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopitheciinae; Cercopithecus.  
 OK NCBI\_TaxID=9531;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RA MEDLINE=93049640; PubMed=1425921;  
 RA Fomsgaard A., Hirsch V.M., Johnson P.R.;  
 RT "Cloning and sequences of primate CD4 molecules: diversity of the  
 RT cellular receptor for simian immunodeficiency virus/human  
 RT immunodeficiency virus.";  
 RL Eur. J. Immunol. 22:2973-2981(1992).  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```

CC -----
DR EMBL: X73328; CAAS1754.1; -.
DR EMBL: X73327; CAAS1753.1; -.
DR HSP: P01730; IMIO.
DR GO: GO:0042101; C:T-cell receptor complex; ISS.
DR GO: GO:0015026; F:coreceptor activity; ISS.
DR GO: GO:0042289; F:MHC class II protein binding; ISS.
DR GO: GO:0006955; P:immune response; ISS.
DR GO: GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO: GO:0030217; P:T-cell differentiation; ISS.
DR GO: GO:0045058; P:T-cell selection; ISS.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig; 2.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PSS0835; IG LIKE; 1.
DR GlycoProfile: Immune_response; Immunoglobulin domain; Lipoprotein;
KM Palmitate; Repeat; T-cell; Transmembrane.
FT NON_TER 1
FT DOMAIN <1 369 Extracellular (Potential).
FT TRANSMEM 370 391 Potential.
FT DOMAIN 392 >397 Cytoplasmic (Potential).
FT DOMAIN <1 98 Ig-like V-type.
FT DOMAIN 99 176 Ig-like C2-type 1.
FT DOMAIN 177 290 Ig-like C2-type 2.
FT DOMAIN 291 347 Ig-like C2-type 3.
FT CARBOHYD 15 15 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 269 269 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 298 298 N-linked (GlcNAc. . .) (Potential).
FT DISULFID 14 82 By similarity.
FT DISULFID 128 157 By similarity.
FT DISULFID 301 343 S-palmitoyl cysteine (By similarity).
FT LIPID 392 392 S-palmitoyl cysteine (By similarity).
FT LIPID 395 395 Missing.
FT VARIANT 20 20 T -> I.
FT VARIANT 43 43 N -> D.
FT VARIANT 86 86 F -> L.
FT VARIANT 96 96 V -> M.
FT VARIANT 173 173 V -> M.
FT VARIANT 316 316 R -> K.
FT NON_TER 397
SQ SEQUENCE 397 AA; 43926 MW; 86608636D2B38A7 CRC64;

Query Match 30.7%; Score 820; DB 1; Length 397;
Best Local Similarity 87.8%; Pred. No. 1.2e-43;
Matches 159; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

```

```

DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Lew-3)
DE (Fragment).
GN Name=CD4.
OS Erythrocybus pataas (Red guenon) (Cercopithecus pataas).
OC Eukaryota; Metazoa; Chordata; Carnata; Vertebrata; Euarchontomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Erythrocybus.
OX NCBI_TaxID=9538;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus."
RL Eur. J. Immunol. 22:2973-2981(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X73324; CAAS1750.1; -.
DR HSP: P01730; IMIO.
DR GO: GO:0042101; C:T-cell receptor complex; ISS.
DR GO: GO:0015026; F:coreceptor activity; ISS.
DR GO: GO:0042289; F:MHC class II protein binding; ISS.
DR GO: GO:0006955; P:immune response; ISS.
DR GO: GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO: GO:0030217; P:T-cell differentiation; ISS.
DR GO: GO:0045058; P:T-cell selection; ISS.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR009973; CD4_TCAG.
DR InterPro: IPR007110; Ig-like.
DR Pfam: PF00047; Ig; 2.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PSS0835; IG LIKE; 1.
DR GlycoProfile: Immune_response; Immunoglobulin domain; Lipoprotein;
KM Palmitate; Repeat; T-cell; Transmembrane.
FT NON_TER 1
FT DOMAIN <1 369 Extracellular (Potential).
FT TRANSMEM 370 391 Potential.
FT DOMAIN 392 >397 Cytoplasmic (Potential).
FT DOMAIN <1 98 Ig-like V-type.
FT DOMAIN 99 176 Ig-like C2-type 1.
FT DOMAIN 177 290 Ig-like C2-type 2.
FT CARBOHYD 269 269 N-linked (GlcNAc. . .) (By similarity).
FT CARBOHYD 298 298 N-linked (GlcNAc. . .) (By similarity).
FT DISULFID 14 82 By similarity.
FT DISULFID 128 157 By similarity.
FT DISULFID 301 343 By similarity.
FT LIPID 392 392 S-palmitoyl cysteine (By similarity).
FT LIPID 395 395 S-palmitoyl cysteine (By similarity).
FT NON_TER 397
SQ SEQUENCE 397 AA; 44081 MW; 67887397A6B7E4F CRC64;

Query Match 30.7%; Score 819; DB 1; Length 397;

```

Best Local Similarity 87.3%; Pred. No. 1.4e-43;  
Matches 158; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

```
QY 28 VVLGKGGDTVELTCTASQKKSIOFHWNKSNQIKILGNQGSFLTGPSSKLNDRADRSRLW 87
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 VVLGKGGDTVELTCTASQKKSIOFHWNKSNQIKILGNQGSFLTGPSSKLNDRADRSRLW 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 88 DQGNPFLIKNKIEDSPYICEVEDQKEVOLVFGLTANSPTHLQOGSLTTLLESP 147
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 DQGCFSMIKNKIEDSEYICEVEDQKEVOLVFGLTANSPTHLQOGSLTTLLESP 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLOQKVEFKIDIVLAFQKA 207
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLOQKVEFKIDIVLAFQKA 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 208 S 208
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 S 181
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

## RESULT 10

```
009261 PRELIMINARY; PRT; 397 AA.
ID 009261;
AC 009261;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE CD4 (Fragment).
GN Name=CD4;
OS Cercopithecus sabaebus.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OC NCBI_TaxID=60711;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98017879; PubMed=9379478;
RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,
RA Corbet S., Barre-Sinoussi F., Allan J.S.;
RA "Relation between phylogeny of African green monkey CD4 genes and
RA their respective simian immunodeficiency virus genes.";
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL; AF001225; AAB60872.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR008424; CD2.
DR InterPro; IPR000973; CD4_TCAG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG v.
DR Pfam; PF05790; C2-sec; Z.
DR Pfam; PF00047; Ig; 1.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR NON_TER 1
FT NON_TER 397
SQ SEQUENCE 397 AA; 43945 MW; 80C963B92A866CD3 CRC64;
```

Query Match 30.1%; Score 803; DB 2; Length 397;  
Best Local Similarity 86.2%; Pred. No. 1.4e-42;  
Matches 156; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

```
QY 28 VVLGKGGDTVELTCTASQKKSIOFHWNKSNQIKILGNQGSFLTGPSSKLNDRADRSRLW 87
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 VVLGKGGDTVELTCTASQKKSIOFHWNKSNQIKILGNQGSFLTGPSSKLNDRADRSRLW 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 88 DQGNPFLIKNKIEDSPYICEVEDQKEVOLVFGLTANSPTHLQOGSLTTLLESP 147
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 DQGCFSMIKNKIEDSEYICEVEDQKEVOLVFGLTANSPTHLQOGSLTTLLESP 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLOQKVEFKIDIVLAFQKA 207
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

Db 121 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLOQKVEFKIDIVLAFQKA 180

QY 208 S 208  
Db 181 S 181

## RESULT 11

```
009262 PRELIMINARY; PRT; 397 AA.
ID 009262;
AC 009262;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE CD4 (Fragment).
GN Name=CD4;
OS Cercopithecus tantalus.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OC NCBI_TaxID=60712;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98017879; PubMed=9379478;
RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,
RA Corbet S., Barre-Sinoussi F., Allan J.S.;
RA "Relation between phylogeny of African green monkey CD4 genes and
RA their respective simian immunodeficiency virus genes.";
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL; AF001221; AAB60868.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR008424; CD2.
DR InterPro; IPR000973; CD4_TCAG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG v.
DR Pfam; PF05790; C2-sec; Z.
DR Pfam; PF00047; Ig; 1.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR NON_TER 1
FT NON_TER 397
SQ SEQUENCE 397 AA; 43954 MW; CFT72P5D82335B0D CRC64;
```

Query Match 30.1%; Score 803; DB 2; Length 397;  
Best Local Similarity 86.2%; Pred. No. 1.4e-42;  
Matches 156; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

```
QY 28 VVLGKGGDTVELTCTASQKKSIOFHWNKSNQIKILGNQGSFLTGPSSKLNDRADRSRLW 87
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 VVLGKGGDTVELTCTASQKKSIOFHWNKSNQIKILGNQGSFLTGPSSKLNDRADRSRLW 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 88 DQGNPFLIKNKIEDSPYICEVEDQKEVOLVFGLTANSPTHLQOGSLTTLLESP 147
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 DQGCFSMIKNKIEDSEYICEVEDQKEVOLVFGLTANSPTHLQOGSLTTLLESP 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLOQKVEFKIDIVLAFQKA 207
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLOQKVEFKIDIVLAFQKA 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 208 S 208
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 S 181
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 12
009263 PRELIMINARY; PRT; 397 AA.
ID 009263;
AC 009263;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
```

```

DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
DE CD4 (Fragment) .
GN Name=CD4;
OS Cercopithecus tantalus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopithecinae; Cercopithecus.
NCBI_TaxID=60712;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98017879; PubMed=9379478;
RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,
RA Corbet S., Barre-Sinoussi F., Allan J.S.;
RT "Relation between phylogeny of African green monkey CD4 genes and
RT their respective simian immunodeficiency virus genes.";
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL: AF001222; AAB60869.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR GO: GO:0006955; P:immune response; IEA.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR00973; CD4_TcAg.
DR InterPro: IPR007110; Ig-Like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF05790; C2-set; 2.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PSS0835; IG_LIKE; 1.
FT NON_TER 1
FT TER 397
SQ SEQUENCE 397 AA; 43994 MW; A3CD031535A51524 CRC64;

Query Match 30.1%; Score 803; DB 2; Length 397;
Best Local Similarity 86.2%; Pred. No. 1.4e-42;
Matches 156; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 28 VILGKGDVIELTCTASQKKSIOFHWKNSNQIKILGNQGSFLTKGPKLNDRADSRSLW 87
DB 1 VILGKGDVIELTCTASQKKSIOFHWKNSNQIKILGNQGSFLTKGPKLNDRADSRSLW 60

QY 88 DQGNFPLIINKLIEDSDTYICEVEDQKEVQLVGLTANSSTHLLQGOSLTLTLESPP 147
DB 61 DQGNFPLIINKLIEDSEYICEVENKKEVELLVGLTANSSTHLLQGOSLTLTLESPP 120

QY 148 GSSPSVQCRSPRGKNIQGGKTLISVSOLEDSGTWTCTVLOKKEVEFKIDIVLAFOKA 207
DB 121 GSSPSVQCRSPRGKNIQGGKTLISVPOLEKDSGTWTCTVSDQNTVEFKIDIVLAFOKA 180

QY 208 S 208
DB 181 S 181

RESULT 13
Q95NE9 PRELIMINARY; PRT; 397 AA.
ID 095NE9;
AC 095NE9;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
DE CD4 (Fragment) .
GN Name=CD4;
OS Cercopithecus pygerythrus (Vervet monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopithecinae; Cercopithecus.
NCBI_TaxID=60710;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98017879; PubMed=9379478;
RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,
RA Corbet S., Barre-Sinoussi F., Allan J.S.;

```

```

RT "Relation between phylogeny of African green monkey CD4 genes and
RT their respective simian immunodeficiency virus genes.";
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL: AF001227; AAB60874.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR GO: GO:0006955; P:immune response; IEA.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR000973; CD4_TcAg.
DR InterPro: IPR007110; Ig-Like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF05790; C2-set; 2.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PSS0835; IG_LIKE; 1.
FT NON_TER 1
FT TER 397
SQ SEQUENCE 397 AA; 43946 MW; 21C3E30882ABFC0 CRC64;

Query Match 30.1%; Score 803; DB 2; Length 397;
Best Local Similarity 86.2%; Pred. No. 1.4e-42;
Matches 156; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 28 VILGKGDVIELTCTASQKKSIOFHWKNSNQIKILGNQGSFLTKGPKLNDRADSRSLW 87
DB 1 VILGKGDVIELTCTASQKKSIOFHWKNSNQIKILGNQGSFLTKGPKLNDRADSRSLW 60

QY 88 DQGNFPLIINKLIEDSDTYICEVEDQKEVQLVGLTANSSTHLLQGOSLTLTLESPP 147
DB 61 DQGNFPLIINKLIEDSEYICEVENKKEVELLVGLTANSSTHLLQGOSLTLTLESPP 120

QY 148 GSSPSVQCRSPRGKNIQGGKTLISVSOLEDSGTWTCTVLOKKEVEFKIDIVLAFOKA 207
DB 121 GSSPSVQCRSPRGKNIQGGKTLISVPOLEKDSGTWTCTVSDQNTVEFKIDIVLAFOKA 180

QY 208 S 208
DB 181 S 181

RESULT 14
O09259 PRELIMINARY; PRT; 397 AA.
ID 009259;
AC 009259;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)
DE CD4 (Fragment) .
GN Name=CD4;
OS Cercopithecus sabaues.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopithecinae; Cercopithecus.
NCBI_TaxID=60711;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98017879; PubMed=9379478;
RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,
RA Corbet S., Barre-Sinoussi F., Allan J.S.;
RT "Relation between phylogeny of African green monkey CD4 genes and
RT their respective simian immunodeficiency virus genes.";
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL: AF001223; AAB60870.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR GO: GO:0006955; P:immune response; IEA.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR000973; CD4_TcAg.
DR InterPro: IPR007110; Ig-Like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF05790; C2-set; 2.

```



DR PRINTS; PR00692; CD4TCANTIGEN.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS50835; IG\_LIKE; 1.  
FT NON\_TER 1 1  
FT NON\_TER 397 397  
SQ SEQUENCE 397 AA; 43881 MW; 7CE39AD0F8506C81 CRC64;

Query Match 29.9%; Score 799; DB 2; Length 397;  
Best Local Similarity 85.6%; Pred. No. 2.5e-42;  
Matches 155; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

QY 28 VVLGKGGDTVELTCAASQKTTTQPHWKNSNQIKILGQGSFLTKGSSKLDRIDSRKSLW 87  
DB 1 VVLGKGGDTVELTCAASQKTTTQPHWKNSNQIKILGQGSFLTKGSSKLDRIDSRKSLW 60  
QY 88 DQGNFPLIIKNIKIEDSDTYICEVEDQKEEVQLVFGLTANSPTHLLOQOSITLTLESPP 147  
DB 61 DQGCFSMIIKNIKIEDSETTYICEVKNKEEVELVFGLTANSPTHLLOQOSITLTLESPP 120  
QY 148 GSSPSVQCRSPRGKNIQSGKTLVSQLELQDSGTWTCVTLQNOKKVEFKIDIVLAFOKA 207  
DB 121 GSSPSVKCRSPRGKNIQSGRTLSVPQLERQDSGTWTCVNSQDQNTVEFKIDIVLAFOKA 180  
QY 208 S 208  
DB 181 S 181

RESULT 15  
009260 PRELIMINARY; PRT; 397 AA.  
AC 009260;  
DT 01-0UL-1997 (TREMBLrel. 04, Created)  
DT 01-0UL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE CD4 (Fragment).  
GN Name=CD4;  
OS Cercopithecus sabaenus.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Cercopithecus.  
OX NCBI\_TaxID=60711;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98017879; PubMed=9379478;  
RA Pomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,  
Corbet S., Barre-Sinoussi F., Allan J.S.;  
RT "Relation between phylogeny of African green monkey CD4 genes and  
their respective simian immunodeficiency virus genes.";  
RU J. Med. Primatol. 26:120-128(1997).  
DR EMBL; AF001224; AB60871.1; -.  
DR GO; GO:0016021; C:Integral to membrane; IEA.  
DR GO; GO:0007155; P:cell adhesion; IEA.  
DR GO; GO:0006955; P:immune response; IEA.  
DR InterPro; IPR008424; CD2.  
DR InterPro; IPR009973; CD4\_TcRg.  
DR InterPro; IPR007110; Ig\_V.  
DR InterPro; IPR003596; Ig\_V.  
DR Pfam; PF05790; C2-set; 2.  
DR Pfam; PF00047; Ig; 1.  
DR PRINTS; PR00692; CD4TCANTIGEN.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS50835; IG\_LIKE; 1.  
FT NON\_TER 1 1  
FT NON\_TER 397 397  
SQ SEQUENCE 397 AA; 43882 MW; 478BB277B92EE89 CRC64;

Query Match 29.9%; Score 799; DB 2; Length 397;  
Best Local Similarity 85.6%; Pred. No. 2.5e-42;  
Matches 155; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

QY 28 VVLGKGGDTVELTCAASQKTTTQPHWKNSNQIKILGQGSFLTKGSSKLDRIDSRKSLW 87  
|||||

DB 1 VVLGKGGDTVELTCAASQKTTTQPHWKNSNQIKILGQGSFLTKGSSKLDRIDSRKSLW 60  
QY 88 DQGNFPLIIKNIKIEDSDTYICEVEDQKEEVQLVFGLTANSPTHLLOQOSITLTLESPP 147  
DB 61 DQGCFSMIIKNIKIEDSETTYICEVKNKEEVELVFGLTANSPTHLLOQOSITLTLESPP 120  
QY 148 GSSPSVQCRSPRGKNIQSGKTLVSQLELQDSGTWTCVTLQNOKKVEFKIDIVLAFOKA 207  
DB 121 GSSPSVKCRSPRGKNIQSGRTLSVPQLERQDSGTWTCVNSQDQNTVEFKIDIVLAFOKA 180  
QY 208 S 208  
DB 181 S 181

Search completed: November 21, 2004, 13:42:25  
Job time : 283.81 secs

**THIS PAGE BLANK (USPTO)**

ORIGIN

Query Match 100.0%; Score 1440; DB 6; Length 1440;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGCTGCACTGGCGCTCTCCCA 60  
 DB 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGCTGCACTGGCGCTCTCCCA 60  
 QY 61 GCAGCCACTCAGAGGAAAGTGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
 DB 61 GCAGCCACTCAGAGGAAAGTGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
 QY 121 TGTACAGCTTCCAGAGAAAGACATCAATTCACCTGGAATAAACTCCACAGATAAG 180  
 DB 121 TGTACAGCTTCCAGAGAAAGACATCAATTCACCTGGAATAAACTCCACAGATAAG 180  
 QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTCCATCCAAAGCTGAATGCGGCT 240  
 DB 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTCCATCCAAAGCTGAATGCGGCT 240  
 QY 241 GACTCAAGAAAGACCTTTGGGACCAAGGAACTTCCCGCTGATCATCAAGAACTTAAG 300  
 DB 241 GACTCAAGAAAGACCTTTGGGACCAAGGAACTTCCCGCTGATCATCAAGAACTTAAG 300  
 QY 301 ATGAAGACTCAGATCTTACATCTGTAAGTGAAGACCAAGAGAGAGAGTCAATTG 360  
 DB 301 ATGAAGACTCAGATCTTACATCTGTAAGTGAAGACCAAGAGAGAGAGTCAATTG 360  
 QY 361 CTAGTGTCCGATGACTGCCAATCTGACACCACTGCTTCAAGGGCAGAGCTTAAC 420  
 DB 361 CTAGTGTCCGATGACTGCCAATCTGACACCACTGCTTCAAGGGCAGAGCTTAAC 420  
 QY 421 CTGACCTTGGAGACCCCGCTGTAAGACCCCTCAGTGCATGTAAGAGTCAAGGGGT 480  
 DB 421 CTGACCTTGGAGACCCCGCTGTAAGACCCCTCAGTGCATGTAAGAGTCAAGGGGT 480  
 QY 481 AAAAACAATACAGGGGGGAGAACCTCTCCGTGTCTGAGCTGAGACTCAGATAGTGC 540  
 DB 481 AAAAACAATACAGGGGGGAGAACCTCTCCGTGTCTGAGCTGAGACTCAGATAGTGC 540  
 QY 541 ACCTGGAATGCACTGTCTTTCAGAACCAAGAAAGTGAAGTTCAAATAGACATGTG 600  
 DB 541 ACCTGGAATGCACTGTCTTTCAGAACCAAGAAAGTGAAGTTCAAATAGACATGTG 600  
 QY 601 GTGCTAGCTTTCAGAGAGGCTCCGAGAGTGGCGGTGTAAGAGGCGGTTCAGGGCGA 660  
 DB 601 GTGCTAGCTTTCAGAGAGGCTCCGAGAGTGGCGGTGTAAGAGGCGGTTCAGGGCGA 660  
 QY 661 GGTGATCCGATGCGAGAGGTTCGGCGGGGTGGAACCGGGGCTGCGCTCCGAGGC 720  
 DB 661 GGTGATCCGATGCGAGAGGTTCGGCGGGGTGGAACCGGGGCTGCGCTCCGAGGC 720  
 QY 721 GGAAGTTCAGAGTGTGCTGAGTCTGGGGCTGAAGTGAAGACCTGGGTCTCTCG 780  
 DB 721 GGAAGTTCAGAGTGTGCTGAGTCTGGGGCTGAAGTGAAGACCTGGGTCTCTCG 780  
 QY 781 GTGAAGGTCTCCGAGAGGCTCTGAGACACCTTCAATCAGATATAGTTTAACTGGAGT 840  
 DB 781 GTGAAGGTCTCCGAGAGGCTCTGAGACACCTTCAATCAGATATAGTTTAACTGGAGT 840  
 QY 841 CGACAGGCCCCCTGAGCAAGGCTTGAAGTGAAGAGATCATCACTATCTTGATGA 900  
 DB 841 CGACAGGCCCCCTGAGCAAGGCTTGAAGTGAAGAGATCATCACTATCTTGATGA 900  
 QY 901 GCACACTACGACCGCACCTTCAGAGGCAAGTCAAGTTACCGGCGCAAGTCCACAGGC 960  
 DB 901 GCACACTACGACCGCACCTTCAGAGGCAAGTCAAGTTACCGGCGCAAGTCCACAGGC 960  
 QY 961 ACAGTCACTGAGAGCGGGAATCTAAGATCGAATCAAGGCGGATATATTTCTGCGG 1020  
 DB 961 ACAGTCACTGAGAGCGGGAATCTAAGATCGAATCAAGGCGGATATATTTCTGCGG 1020  
 QY 1021 GGAAGTACGAGAGGAGCGGACGAGAGGGGAATATGATATATATGAGTTTCTGAACAT 1080

DB 1021 GGAAGTACGAGAGGAGCGGACGAGAGGGGAATATGATATATATGAGTTTCTGAACAT 1080  
 QY 1081 TGGGGCCAGGAAACCTTGCTCAGGTCACCTCAAGTGGCGGTGGCTCCGAGGTGGTGG 1140  
 DB 1081 TGGGGCCAGGAAACCTTGCTCAGGTCACCTCAAGTGGCGGTGGCTCCGAGGTGGTGG 1140  
 QY 1141 AGCGGTGGCGGGAATCTGAATCTGAGTTGACCGAGTCTCAGCCACCTGTCTGTCT 1200  
 DB 1141 AGCGGTGGCGGGAATCTGAATCTGAGTTGACCGAGTCTCAGCCACCTGTCTGTCT 1200  
 QY 1201 CCAAGGAAAGACCACTCTCTCTGAGGCGCAGTGAAGTGTATGACGACTTAAGCC 1260  
 DB 1201 CCAAGGAAAGACCACTCTCTCTGAGGCGCAGTGAAGTGTATGACGACTTAAGCC 1260  
 QY 1261 TGGTACCAAGAAACCTGGCCAGGCTCCAGGCTCTCTCATATATGTGTATCCACAGG 1320  
 DB 1261 TGGTACCAAGAAACCTGGCCAGGCTCCAGGCTCTCTCATATATGTGTATCCACAGG 1320  
 QY 1321 GCCACCGGTGCCAGCAGGTCAGTGGCAGTGGGTCTGGGCAAAATCACTTCACC 1380  
 DB 1321 GCCACCGGTGCCAGCAGGTCAGTGGCAGTGGGTCTGGGCAAAATCACTTCACC 1380  
 QY 1381 ATCAGACCTGAGTCTGAAGATTTTGCAGTTTATCTGTACAGATCAATTAACCTGG 1440  
 DB 1381 ATCAGACCTGAGTCTGAAGATTTTGCAGTTTATCTGTACAGATCAATTAACCTGG 1440

RESULT 2  
 AR363716 1213 bp DNA linear PAT 03-SEP-2003  
 LOCUS AR363716  
 DEFINITION Sequence 1 from patent US 5223418.  
 ACCESSION AR363716  
 VERSION AR363716.1 GI:34425655  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 1213)  
 AUTHORS Arcuri,E.J., Brawner,M.E., Donovan,M.J., Gerber,R.G. and Keller,J.A.  
 TITLE Method of Improving the yield of heterologous proteins produced by Streptomyces lividans  
 JOURNAL Patent: US 5223418-A 1 29-JUN-1993;  
 FEATURES  
 source 1..1213  
 /organism="unknown"  
 /mol\_type="genomic DNA"

ORIGIN  
 Query Match 43.4%; Score 625; DB 6; Length 1213;  
 Best Local Similarity 97.7%; Pred. No. 1.3e-134;  
 Matches 634; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGCTGCACTGGCGCTCTCCCA 60  
 DB 79 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGCTGCACTGGCGCTCTCCCA 138  
 QY 61 GCAGCCACTCAGAGGAAAGTGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
 DB 139 GCAGCCACTCAGAGGAAAGTGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 198  
 QY 121 TGTACAGCTTCCAGAGAAAGACATCAATTCACCTGGAATAAACTCCACAGATAAG 180  
 DB 199 TGTACAGCTTCCAGAGAAAGACATCAATTCACCTGGAATAAACTCCACAGATAAG 258  
 QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTCCATCCAAAGCTGAATGATGCGGCT 240  
 DB 259 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTCCATCCAAAGCTGAATGATGCGGCT 318  
 QY 241 GACTCAAGAAAGACCTTTGGGACCAAGGAACTTCCCGCTGATCATCAAGAAATCTTAAG 300  
 DB 319 GACTCAAGAAAGACCTTTGGGACCAAGGAACTTCCCGCTGATCATCAAGAAATCTTAAG 378

QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGGAGGAGGTCGCAATTG 360  
DB 379 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGGAGGAGGTCGCAATTG 438  
QY 361 CTAGTGTTCGATTTGATCTGCACTCTGACACCCACTCTCTTCAAGGGGAGAGCCTGACC 420  
DB 439 CTAGTGTTCGATTTGATCTGCACTCTGACACCCACTCTCTTCAAGGGGAGAGCCTGACC 498  
QY 421 CTGACCTTTGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATGTAGAGTCCAGAGGGGT 480  
DB 499 CTGACCTTTGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATGTAGAGTCCAGAGGGGT 558  
QY 481 AAAAATCATACAGAGGGGAGAGAGCCCTCTCGTGTCTCAGTGCATGTAGAGTCCAGAGGGGT 540  
DB 559 AAAAATCATACAGAGGGGAGAGAGCCCTCTCGTGTCTCAGTGCATGTAGAGTCCAGAGGGGT 618  
QY 541 ACCTGACATGACACTGTCTTTGAGAGACCAAGAGAGTGAAGTCAAAATAGACATCTGTG 600  
DB 619 ACCTGACATGACACTGTCTTTGAGAGACCAAGAGAGTGAAGTCAAAATAGACATCTGTG 678  
QY 601 GTGCTAGCTTTCCAGAGAGCCTCTCCGAGGTGGCGGTAGTGGGGAGGCG 649  
DB 679 GTGCTAGCTTTCCAGAGAGCCTCTCCGAGGTGGCGGTAGTGGGGAGGCG 727

RESULT 3  
AX100880 1377 bp DNA linear PAT 10-APR-2001  
LOCUS Sequence 30 from Patent WO0122084.  
DEFINITION AX100880  
ACCESSION AX100880  
VERSION AX100880.1 GI:13619792  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1  
AUTHORS Jakobson, B.K.  
TITLE Spr identification of inhibitors of receptor-ligand interactions  
JOURNAL Patent: WO 0122084-A 30 29-MAR-2001;  
Avidex Ltd (GB)

FEATURES  
source Location/Qualifiers

## CDS

1..1377  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
1..1377  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC36495.1"  
/db\_xref="GI:13619793"  
/translation="MNRGVPRFRLILVQLALLPATQSKVVLKGGPVTVELTCTAS  
OKSIOFHKNSNNOIKLNGOSFLTKGSKLNDASRSKLDGNGPLIKNLKIE  
DSDTICEVEDQKEVOLIVPLTANSIDHILGOSLITLSPGSSISYCKESDPRG  
KNIGGKTLVSQLELDQSGTTCVYLNOQKVEKIDIVLAFQKASISYKKEGEO  
VEFSPLAFVYKLTGSGELMWQARASSSKSWITFDLKNKSVKRVATQDPKIQMGK  
KLPLHLTLPQALPOYAGSGNLTLALEANTGKLHDEVNLVVRATQLOKNLCEWGP  
SPKMLSLKLENKEAKVSRKRAVAVLNPEDAMQCLSDSGVLLSNIKVLPWST  
FVQPMALIVLGVAGLILFLIGLIFPCVRCRRRRQAEBSQIKRLLSBKTKCQCPHR  
FOKTCSPPI"

## ORIGIN

Query Match 43.4%; Score 624.4; DB 6; Length 1377;  
Best Local Similarity 99.8%; Pred. No. 1.7e-134;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTGTGTGCTGCAACTGGCGCTCTCCCA 60  
DB 1 ATGAACCGGGAGTCCCTTTAGGCACTTGTGTGCTGCTCAACTGGCGCTCTCCCA 60  
QY 61 GCAGCACTCAGGAGAAAGAGTGTGCTGGCAAAAAGGGGATACAGTGAACGTACC 120

DB 61 GCAGCACTCAGGAGAAAGAGTGTGCTGGCAAAAAGGGGATACAGTGAACGTACC 120  
QY 121 TGTACAGCTTCCCAAGAGAGCATACATTTCCACTGGAAAACTCCACAGATTAAG 180  
DB 121 TGTACAGCTTCCCAAGAGAGCATACATTTCCACTGGAAAACTCCACAGATTAAG 180  
QY 181 ATTCTGGGAAATCAGGGGCTCTTTAACTAAAGGTTCATCCAGCTGAATGATGCGGT 240  
DB 181 ATTCTGGGAAATCAGGGGCTCTTTAACTAAAGGTTCATCCAGCTGAATGATGCGGT 240  
QY 241 GACTAAGAGAGAGCTTTGGAGCCAGAGAACTTCCCTGTATCATCAAGATCTTAAG 300  
DB 241 GACTAAGAGAGAGCTTTGGAGCCAGAGAACTTCCCTGTATCATCAAGATCTTAAG 300  
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGGAGTCAATTG 360  
DB 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGGAGTCAATTG 360  
QY 361 CTAGTGTTCGATTTGATCTGCACTCTGACACCCACTCTCTTCAAGGGGAGAGCCTGACC 420  
DB 361 CTAGTGTTCGATTTGATCTGCACTCTGACACCCACTCTCTTCAAGGGGAGAGCCTGACC 420  
QY 421 CTGACCTTTGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATGTAGAGTCCAGAGGGGT 480  
DB 421 CTGACCTTTGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATGTAGAGTCCAGAGGGGT 480  
QY 481 AAAAATCATACAGAGGGGAGAGAGCCCTCTCGTGTCTCAGTGCATGTAGAGTCCAGAGGGGT 540  
DB 481 AAAAATCATACAGAGGGGAGAGAGCCCTCTCGTGTCTCAGTGCATGTAGAGTCCAGAGGGGT 540  
QY 541 ACCTGACATGACACTGTCTTTGAGAGACCAAGAGAGTGAAGTCAAAATAGACATCTGTG 600  
DB 541 ACCTGACATGACACTGTCTTTGAGAGACCAAGAGAGTGAAGTCAAAATAGACATCTGTG 600  
QY 601 GTGCTAGCTTTCCAGAGAGCCTCTCCGAGGTGGCGGTAGTGGGGAGGCG 649  
DB 601 GTGCTAGCTTTCCAGAGAGCCTCTCCGAGGTGGCGGTAGTGGGGAGGCG 649

RESULT 4  
AR380468 1742 bp DNA linear PAT 18-DEC-2003  
LOCUS Sequence 1013 from patent US 6607879.  
DEFINITION AR380468  
ACCESSION AR380468  
VERSION AR380468.1 GI:40088102  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1742)  
AUTHORS Cocks, B.G., Stuart, S.G. and Seilhamer, J.J.  
TITLE Compositions for the detection of blood cell and immunological  
response gene expression  
JOURNAL Patent: US 6607879-A 1013 19-AUG-2003;  
FEATURES  
source Location/Qualifiers

## ORIGIN

Query Match 43.4%; Score 624.4; DB 6; Length 1742;  
Best Local Similarity 99.8%; Pred. No. 1.7e-134;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTGTGTGCTGCAACTGGCGCTCTCCCA 60  
DB 76 ATGAACCGGGAGTCCCTTTAGGCACTTGTGTGCTGCTCAACTGGCGCTCTCCCA 135  
QY 61 GCAGCACTCAGGAGAAAGAGTGTGCTGGCAAAAAGGGGATACAGTGAACGTACC 120  
DB 136 GCAGCACTCAGGAGAAAGAGTGTGCTGGCAAAAAGGGGATACAGTGAACGTACC 195  
QY 121 TGTACAGCTTCCCAAGAGAGCATACATTTCCACTGGAAAACTCCACAGATTAAG 180

|||||  
Db 196 TGTACAGCTTCCAGAGAAAGACATACATTCACCTGGAAGAACTCCACCAAGATTAAG 255  
Qy 181 ATTCGGGAATCAGGGCTCCTTCTTAACCTAAGCTCATCCAGCTGAAGATCGGCT 240  
Db 256 ATTCGGGAATCAGGGCTCCTTCTTAACCTAAGCTCATCCAGCTGAAGATCGGCT 315  
Qy 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG 300  
Db 316 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG 375  
Qy 301 ATGAGAGCTCAGATTAATTACTTGTGAGGTGAGACCAAGAGAGAGAGTCAATTG 360  
Db 376 ATGAGAGCTCAGATTAATTACTTGTGAGGTGAGACCAAGAGAGAGAGTCAATTG 435  
Qy 361 CTAGTGTTCGGATTGACTGCACTGCAACCACTGCTTCAAGGGCAGAGCTGACC 420  
Db 436 CTAGTGTTCGGATTGACTGCACTGCAACCACTGCTTCAAGGGCAGAGCTGACC 495  
Qy 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGAATGTAGAGTCCAAGGGT 480  
Db 496 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGAATGTAGAGTCCAAGGGT 555  
Qy 481 AAAAACAATACAGGGGGGAAAGACCTTCTCGTGTCTCAGCTGAGCTCCAGATATGTC 540  
Db 556 AAAAACAATACAGGGGGGAAAGACCTTCTCGTGTCTCAGCTGAGCTCCAGATATGTC 615  
Qy 541 ACCTGGACATGCACTGTCTTGGACAAACCAAGAAAGGTGAGTTCAAAATAGACATGTG 600  
Db 616 ACCTGGACATGCACTGTCTTGGACAAACCAAGAAAGGTGAGTTCAAAATAGACATGTG 675  
Qy 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
Db 676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701

RESULT 5  
AX287109 1742 bp DNA linear PAT 21-NOV-2001  
LOCUS AX287109  
DEFINITION Sequence 7 from Patent WO0164752.  
ACCESSION AX287109  
VERSION AX287109.1 GI:17049085  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Liltman,D.R., Kwon,D., van Kooyk,Y.C. and Geijtenbeek,T.C.  
TITLE Antibody inhibiting the binding between gp120 and dc-si3n and  
screening methods  
JOURNAL Patent: WO 0164752-A 7 07-SEP-2001;  
NEW YORK UNIVERSITY (US); KATHOLIEKE UNIVERSITEIT NIJMEGEN (NL)  
FEATURES  
source  
1. 1742  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 43.4%; Score 624.4; DB 6; Length 1742;  
Best Local Similarity 99.8%; Pred. No. 1.7e-134;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 ATGAACCGGGAGTCCCTTTAGGCACTTCTTGGTGTGCAACTGGCCTCTCCCA 60  
Db 76 ATGAACCGGGAGTCCCTTTAGGCACTTCTTGGTGTGCAACTGGCCTCTCCCA 135  
Qy 61 GCACCACTCAGAGAAAGAGTGTGTGGCAAAAAAGGGGATACAGTGAATGACC 120  
Db 136 GCACCACTCAGAGAAAGAGTGTGTGGCAAAAAAGGGGATACAGTGAATGACC 195  
Qy 121 TGTACAGCTTCCAGAGAAAGACATACATTCACCTGGAAGAACTCCACCAAGATTAAG 180

|||||  
Db 196 TGTACAGCTTCCAGAGAAAGACATACATTCACCTGGAAGAACTCCACCAAGATTAAG 255  
Qy 181 ATTCGGGAATCAGGGCTCCTTCTTAACCTAAGCTCATCCAGCTGAAGATCGGCT 240  
Db 256 ATTCGGGAATCAGGGCTCCTTCTTAACCTAAGCTCATCCAGCTGAAGATCGGCT 315  
Qy 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG 300  
Db 316 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG 375  
Qy 301 ATGAGAGCTCAGATTAATTACTTGTGAGGTGAGACCAAGAGAGAGAGTCAATTG 360  
Db 376 ATGAGAGCTCAGATTAATTACTTGTGAGGTGAGACCAAGAGAGAGAGTCAATTG 435  
Qy 361 CTAGTGTTCGGATTGACTGCACTGCAACCACTGCTTCAAGGGCAGAGCTGACC 420  
Db 436 CTAGTGTTCGGATTGACTGCACTGCAACCACTGCTTCAAGGGCAGAGCTGACC 495  
Qy 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGAATGTAGAGTCCAAGGGT 480  
Db 496 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGAATGTAGAGTCCAAGGGT 555  
Qy 481 AAAAACAATACAGGGGGGAAAGACCTTCTCGTGTCTCAGCTGAGCTCCAGATATGTC 540  
Db 556 AAAAACAATACAGGGGGGAAAGACCTTCTCGTGTCTCAGCTGAGCTCCAGATATGTC 615  
Qy 541 ACCTGGACATGCACTGTCTTGGACAAACCAAGAAAGGTGAGTTCAAAATAGACATGTG 600  
Db 616 ACCTGGACATGCACTGTCTTGGACAAACCAAGAAAGGTGAGTTCAAAATAGACATGTG 675  
Qy 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
Db 676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701

RESULT 6  
HUMATCT4 1742 bp mRNA linear PRI 27-APR-1993  
LOCUS HUMATCT4  
DEFINITION Human T-cell surface glycoprotein T4 mRNA, complete cde.  
ACCESSION M12807  
VERSION M12807.1 GI:179141  
KEYWORDS  
glycoprotein; immunoglobulin super gene family; recognition  
antigen; surface antigen.  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 1742)  
Maddon,P.J., Liltman,D.R., Godfrey,M., Maddon,D.E., Chess,L. and  
Axel,R.  
TITLE The isolation and nucleotide sequence of a cDNA encoding the T cell  
surface protein T4; a new member of the immunoglobulin gene family  
JOURNAL Cell 42 (1), 93-104 (1985)  
MEDLINE 85254948  
PUBMED 2990730  
REFERENCE 2 (bases 153 to 153)  
Liltman,D.R.  
AUTHORS  
JOURNAL Unpublished (1986)  
COMMENT  
Original source text: Human peripheral T lymphocyte, cDNA to mRNA,  
clone p14B.  
T4 is a member of the immunoglobulin supergene family. The T4 mRNA  
encodes contiguous V- and J-like elements without the requirement  
for DNA recombination events. The V-like domain can be found at  
positions 145 to 426 and the J-like domain from 427 to 471. The T4  
protein may serve as the specific surface receptor for the AIDS  
virus [1].  
With the revision of the nucleotide at position 153, the amino acid  
encoded by 151-153 becomes lysine. This is now regarded to be the  
first residue in the mature protein [2].  
Location/Qualifiers  
1. 1742  
/organism="Homo sapiens"

FEATURES  
source

CDS

/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
76..1452  
/note="T4 surface glycoprotein precursor"  
/codon\_start=1  
/protein\_id="AA35572.1"  
/db\_xref="GI:179142"  
/translation="MNRGVPPRHLLVLQALLPAATOGKVVLAGKGDVETLCTAS  
OKSIQPHWKNNOIKILGNOSFLTKGSKLNDPDRSLMDGNPFLIKNLKIE  
DSDDYICEVDEKREYQLVFGILTASDTHLQOGSLTLLSPSSSPVOCSPRG  
KNIQGGKTLVSQLELDSDGTWTCVLQNKVEFKIDIVLAFKASSIVYKKEBO  
VFSPPLATVEKLTGSGELMWQERASSKSWTFDLKNEVSVKWTQDPLQMGK  
KLPHLTLPQALPOYAGSGLNLTALAEATGKHQEVNLVVRATDLOKRLCEWGP  
SPKMLSLKLNKKAIVSKREKAVWLNPEAGMOCILSDSGOVLLESNIKVLPWST  
PVOPMALIVLGVAGLLFLFIGLGFPCVCRHRRRQARMNSQIKRLSEKKTCCGPHR  
FQKTCSP"

sig\_peptide  
76..150  
/note="T4 surface glycoprotein signal peptide"  
mat\_peptide  
151..1449  
/product="T4 surface glycoprotein mature peptide"

ORIGIN 198 bp upstream of RsaI site.

Query Match 43.4%; Score 624.4; DB 9; Length 1742;  
Best Local Similarity 99.8%; Pred. No. 1.7e-134;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAACCGGGAGTCCCTTTAGGCACTTGCTGTGCTGCAACTGGCGCTCCCA 60  
76 ATGAACCGGGAGTCCCTTTAGGCACTTGCTGTGCTGCTGCAACTGGCGCTCCCA 135  
61 GCAGCACTCAGGGAAGAAAGTGTGTGCGCAAAAAGGGGATACAGTGAACCTGACC 120  
136 GCAGCACTCAGGGAAGAAAGTGTGTGCGCAAAAAGGGGATACAGTGAACCTGACC 195  
121 TGTACAGCTTCCAGGAAGAGCATACATTCACCTGGAATACTCCACCAAGATAAG 180  
196 TGTACAGCTTCCAGGAAGAGCATACATTCACCTGGAATACTCCACCAAGATAAG 255  
101 ATTCGGGAATCAGAGGCTCTTCTTAATTAAGTCCATCCAGTGAATGATGCGCT 240  
256 ATTCGGGAATCAGAGGCTCTTCTTAATTAAGTCCATCCAGTGAATGATGCGCT 315  
241 GACTCAGGAAGAGCTTTGGAGCAAGAACTCCCGCTGATCATCAAGATCTTAG 300  
316 GACTCAGGAAGAGCTTTGGAGCAAGAACTCCCGCTGATCATCAAGATCTTAG 375  
301 ATGAAGACTCAGATTAATTATCTGTGAAGTGAAGACAGAAAGGAGGTGCAATTG 360  
376 ATGAAGACTCAGATTAATTATCTGTGAAGTGAAGACAGAAAGGAGGTGCAATTG 435  
361 CTAGTGTTCGATTAATCTGTGACCTGACACCTGCTTCAGGGGCAAGCTGACC 420  
436 CTAGTGTTCGATTAATCTGTGACCTGACACCTGCTTCAGGGGCAAGCTGACC 495  
421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGAATGTAAGAGTCCAGGGGT 480  
496 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGAATGTAAGAGTCCAGGGGT 555  
481 AAAAATCAGAGAGGGGGGAGAACCTCTTCCTGTGTCTCAGCTGAGCTCCAGATAGTGC 540  
556 AAAAATCAGAGAGGGGGGAGAACCTCTTCCTGTGTCTCAGCTGAGCTCCAGATAGTGC 615  
541 ACCTGGAATGACATCTCTTGTGAAACCAAGAAAGGTGAGATTCAAAATGACATCTGTG 600  
616 ACCTGGAATGACATCTCTTGTGAAACCAAGAAAGGTGAGATTCAAAATGACATCTGTG 675  
601 GTGTAGCTTTCCAGAAAGGCTCCCG 626  
676 GTGTAGCTTTCCAGAAAGGCTCCCG 701

LOCUS HUMACT4A 1742 bp DNA linear PRI 24-JUN-1994  
DEFINITION Human T4 surface glycoprotein CD4 gene, complete cds.  
ACCESSION M35160.1  
VERSION M35160.1 GI:179143  
KEYWORDS glycoprotein; immunoglobulin super gene family; recognition antigen; surface antigen.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 1742)  
Hodge, T.W., Sasse, D.R. and McDougal, J.S.  
Humans with OKT4-epitope deficiency have a single nucleotide base change in the CD4 gene, resulting in substitution of TRP240 for ARG240  
JOURNAL Hum. Immunol. 30 (2), 99-104 (1991)  
MEDLINE 91216786  
PUBMED 1708753  
REFERENCE 2 (bases 1 to 1742)  
Hodge, T.W.  
Direct Submission  
Submitted (12-JUN-1990) T.W. Hodge, A-25 Bldg. 1, Rm. 1226, Centers for Disease Control, 1600 Clifton Rd. Atlanta, GA 30333 USA  
COMMENT Original source text: Human DNA.  
A cytosine to thymidine transition at nucleotide position 868 results in substitution of TRP-240 for ARG-240.  
location/Qualifiers  
1..1742  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/haploidy="OKT4-negative"  
76..1452  
/gene="CD4"  
76..1452  
/note="CD4"  
/product="T4 surface glycoprotein precursor"  
/codon\_start=1  
/protein\_id="AA16069.1"  
/db\_xref="GI:179144"  
/translation="MNRGVPPRHLLVLQALLPAATOGKVVLAGKGDVETLCTAS  
OKSIQPHWKNNOIKILGNOSFLTKGSKLNDPDRSLMDGNPFLIKNLKIE  
DSDDYICEVDEKREYQLVFGILTASDTHLQOGSLTLLSPSSSPVOCSPRG  
KNIQGGKTLVSQLELDSDGTWTCVLQNKVEFKIDIVLAFKASSIVYKKEBO  
VFSPPLATVEKLTGSGELMWQERASSKSWTFDLKNEVSVKWTQDPLQMGK  
KLPHLTLPQALPOYAGSGLNLTALAEATGKHQEVNLVVRATDLOKRLCEWGP  
SPKMLSLKLNKKAIVSKREKAVWLNPEAGMOCILSDSGOVLLESNIKVLPWST  
PVOPMALIVLGVAGLLFLFIGLGFPCVCRHRRRQARMNSQIKRLSEKKTCCGPHR  
FQKTCSP"

sig\_peptide  
76..150  
/gene="CD4"  
/note="T4 surface glycoprotein signal peptide"  
mat\_peptide  
151..1449  
/gene="CD4"  
/product="T4 surface glycoprotein mature peptide"  
variation  
792  
/gene="CD4"  
/note="g in OKT4 epitope deficiency; c in wt"

ORIGIN

Query Match 43.4%; Score 624.4; DB 9; Length 1742;  
Best Local Similarity 99.8%; Pred. No. 1.7e-134;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAACCGGGAGTCCCTTTAGGCACTTGCTGTGCTGCAACTGGCGCTCCCA 60  
76 ATGAACCGGGAGTCCCTTTAGGCACTTGCTGTGCTGCTGCAACTGGCGCTCCCA 135  
61 GCAGCACTCAGGGAAGAAAGTGTGTGCGCAAAAAGGGGATACAGTGAACCTGACC 120  
136 GCAGCACTCAGGGAAGAAAGTGTGTGCGCAAAAAGGGGATACAGTGAACCTGACC 195  
121 TGTACAGCTTCCAGGAAGAGCATACATTCACCTGGAATACTCCACCAAGATAAG 180

|||||  
Db 196 TGTACAGCTTCCCGAAGAGATCAATTCACGTGAAAAAACTCCACCAAGATAAAG 255  
|||  
Qy 181 ATTCTGGGAATCAGGGCTCCTTTTAACATAAGTCCATCCAGCTGAATGATCGGCT 240  
Db 256 ATTCTGGGAATCAGGGCTCCTTTTAACATAAGTCCATCCAGCTGAATGATCGGCT 315  
|||  
Qy 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 300  
Db 316 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 375  
|||  
Qy 301 ATAGAACTCAGATTAATCAATCTGTGAAGTGAAGACCAAGAAAGAGAGTCAATTG 360  
Db 376 ATAGAACTCAGATTAATCAATCTGTGAAGTGAAGACCAAGAAAGAGAGTCAATTG 435  
|||  
Qy 361 CTAGTGTTCGATTTGACTGTCCAACTGACACCCACTGCTTCAGGGGCGAGGCTACC 420  
Db 436 CTAGTGTTCGATTTGACTGTCCAACTGACACCCACTGCTTCAGGGGCGAGGCTACC 495  
|||  
Qy 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGCAATGTAGAGATCCAAAGGGT 480  
Db 496 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGCAATGTAGAGATCCAAAGGGT 555  
|||  
Qy 481 AAAAACAATACAGGGGGGAAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGC 540  
Db 556 AAAAACAATACAGGGGGGAAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGC 615  
|||  
Qy 541 ACCTGACATGCACTGTCTTGGCAGAACCAAGAAAGGTGAGTTCAAAATAGACATGTG 600  
Db 616 ACCTGACATGCACTGTCTTGGCAGAACCAAGAAAGGTGAGTTCAAAATAGACATGTG 675  
|||  
Qy 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626  
Db 676 GTGCTAGCTTTCAGAAAGGCTCCAG 701

RESULT 8  
LOCUS 108116 1910 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 3 from Patent EP 0325262.  
ACCESSION 108116  
VERSION 108116.1 GI:589172  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1910)  
AUTHORS Brian,S.D.  
TITLE Cloned genes encoding IG-CD4 fusion proteins and the use thereof  
JOURNAL Patent: EP 0325262-A2 3 26-JUL-1989;  
FEATURES  
source location/Qualifiers  
1..1910  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 43.4%; Score 624.4; DB 6; Length 1910;  
Best Local Similarity 99.8%; Pred. No. 1.7e-134;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
|||  
Qy 1 ATGAACCGGGAGTCCCTTTTAGCACTTGCTTGCTGTGCTGAACCTGCGCTCTCCCA 60  
Db 111 ATGAACCGGGAGTCCCTTTTAGCACTTGCTTGCTGTGCTGAACCTGCGCTCTCCCA 170  
|||  
Qy 61 GCAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAATGACC 120  
Db 171 GCAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAATGACC 230  
|||  
Qy 121 TGTACAGCTTCCCAAGAAAGAGATCAATTCACAGTGAAGAACTCCACAGATAAAG 180  
Db 231 TGTACAGCTTCCCAAGAAAGAGATCAATTCACAGTGAAGAACTCCACAGATAAAG 290  
|||  
Qy 181 ATTCTGGGAATCAGGGCTCCTTTTAACATAAGTCCATCCAGCTGAATGATCGGCT 240

|||||  
Db 291 ATTCTGGGAATCAGGGCTCCTTTTAACATAAGTCCATCCAGCTGAATGATCGGCT 350  
|||  
Qy 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 300  
Db 351 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 410  
|||  
Qy 301 ATAGAACTCAGATTAATCAATCTGTGAAGTGAAGACCAAGAAAGAGAGTCAATTG 360  
Db 411 ATAGAACTCAGATTAATCAATCTGTGAAGTGAAGACCAAGAAAGAGAGTCAATTG 470  
|||  
Qy 361 CTAGTGTTCGATTTGACTGTCCAACTGACACCCACTGCTTCAGGGGCGAGGCTACC 420  
Db 471 CTAGTGTTCGATTTGACTGTCCAACTGACACCCACTGCTTCAGGGGCGAGGCTACC 530  
|||  
Qy 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGCAATGTAGAGATCCAAAGGGT 480  
Db 531 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGCAATGTAGAGATCCAAAGGGT 590  
|||  
Qy 481 AAAAACAATACAGGGGGGAAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGC 540  
Db 591 AAAAACAATACAGGGGGGAAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGC 650  
|||  
Qy 541 ACCTGACATGCACTGTCTTGGCAGAACCAAGAAAGGTGAGTTCAAAATAGACATGTG 600  
Db 651 ACCTGACATGCACTGTCTTGGCAGAACCAAGAAAGGTGAGTTCAAAATAGACATGTG 710  
|||  
Qy 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626  
Db 711 GTGCTAGCTTTCAGAAAGGCTCCAG 736

RESULT 9  
LOCUS 108115 3133 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent EP 0325262.  
ACCESSION 108115  
VERSION 108115.1 GI:589171  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 3133)  
AUTHORS Brian,S.D.  
TITLE Cloned genes encoding IG-CD4 fusion proteins and the use thereof  
JOURNAL Patent: EP 0325262-A2 1 26-JUL-1989;  
FEATURES  
source location/Qualifiers  
1..3133  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 43.4%; Score 624.4; DB 6; Length 3133;  
Best Local Similarity 99.8%; Pred. No. 1.7e-134;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
|||  
Qy 1 ATGAACCGGGAGTCCCTTTTAGCACTTGCTTGCTGTGCTGAACCTGCGCTCTCCCA 60  
Db 111 ATGAACCGGGAGTCCCTTTTAGCACTTGCTTGCTGTGCTGAACCTGCGCTCTCCCA 170  
|||  
Qy 61 GCAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAATGACC 120  
Db 171 GCAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAATGACC 230  
|||  
Qy 121 TGTACAGCTTCCCAAGAAAGAGATCAATTCACAGTGAAGAACTCCACAGATAAAG 180  
Db 231 TGTACAGCTTCCCAAGAAAGAGATCAATTCACAGTGAAGAACTCCACAGATAAAG 290  
|||  
Qy 181 ATTCTGGGAATCAGGGCTCCTTTTAACATAAGTCCATCCAGCTGAATGATCGGCT 240  
Db 291 ATTCTGGGAATCAGGGCTCCTTTTAACATAAGTCCATCCAGCTGAATGATCGGCT 350  
|||  
Qy 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 300



```
Db 351 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 410
Qy 301 ATAGAAGACTCATGATCTTACATCTGTGAAGTGAGAGACCAAGAGAGAGTGCAATTG 360
Db 411 ATAGAAGACTCATGATCTTACATCTGTGAAGTGAGAGACCAAGAGAGAGTGCAATTG 470
Qy 361 CTAGTGTGGATTGATCTGCACTGTGACACCCACTGCTTACAGGGGCAAGACCTGACC 420
Db 471 CTAGTGTGGATTGATCTGCACTGTGACACCCACTGCTTACAGGGGCAAGACCTGACC 530
Qy 421 CTGACCTTGGAGAGCCCCCTGGTAGTACCCCTCAGTCAATGTAGAGTCCAGAGGGGT 480
Db 531 CTGACCTTGGAGAGCCCCCTGGTAGTACCCCTCAGTCAATGTAGAGTCCAGAGGGGT 590
Qy 481 AAAAACAATCAGGGGGGGAAGACCTCTCTCTGCTGCTGCTGAGTGTCCAGATAGTGGC 540
Db 591 AAAAACAATCAGGGGGGGAAGACCTCTCTCTGCTGCTGCTGAGTGTCCAGATAGTGGC 650
Qy 541 ACCTGAGCATGCACTGCTTGTGAGACCAAGAGAGTGAAGTCAAAATAGACATCGTG 600
Db 651 ACCTGAGCATGCACTGCTTGTGAGACCAAGAGAGTGAAGTCAAAATAGACATCGTG 710
Qy 601 GTGCTAGCTTTCCAGAGGCGCTCCGG 626
Db 711 GTGCTAGCTTTCCAGAGGCGCTCCAG 736

RESULT 10
AY438650 7533 bp DNA linear SYN 11-NOV-2003
LOCUS Cloning vector pMCS1Bac, complete sequence.
DEFINITION AY438650
ACCESSION AY438650.1 GI:38197734
VERSION
KEYWORDS
SOURCE Cloning vector pMCS1Bac
ORGANISM Cloning vector pMCS1Bac
artificial sequences; vectors.
REFERENCE 1 (bases 1 to 7533)
AUTHORS Phillips,B., Foretner,M. and Mayr,L.M.
TITLE A baculovirus expression system for magnetic sorting of infected
cells and enhanced titer determination
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 7533)
AUTHORS Phillips,B., Foretner,M. and Mayr,L.M.
TITLE Direct Subinjection
JOURNAL Submitted (16-OCT-2003) IDC / BSO1, Novartis Pharma AG, WSJ-88.601,
Basel CH-4002, Switzerland
FEATURES
SOURCE Location/Qualifiers
1.7533
/organism="Cloning vector pMCS1Bac"
/mol_type="other DNA"
/db_xref="taxon:254251"
689..1549
/gene="bia"
689..1549
/gene="bia"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AAR13800.1"
/db_xref="GI:38197735"
/transl_table=11
/translation="MSIQHFRVALIPFPAFLPVPAPHEPILVKVKADQDQAGVGY
LELDNSGKILSFPEERFPFPMSTPKVILCGAVSRIDAGQSGRRIHYSQNDLVE
YSPVTEKHLIDMTGTRRELCSAAITMSDNTAAILLTITGGPEKELTAPFLHNMGDHTRL
DRMEPELNBAIPNDERDITMPVAAATTAKLITGELLTLASQOOLIDMWEADKVAQPL
LRSLAPAGWFIADKSGAGBGRSGIILALGPGRSRIYIYITTSQATMBERNQIA
EIGASLIKHW"
complement (2902..3435)
/gene="aacc1"
complement (2902..3435)
/gene="aacc1"
/codon_start=1
```

```
/transl_table=11
/product="gentamicin-3-acetyl transferase"
/protein_id="AAR13801.1"
/db_xref="GI:38197736"
/transl_table=11
/translation="MLRSSNDVTQOGRSPKTKLGSSMGIIIFTCRLGPDQVSKMBAL
DLFGEFPGVATYSOHOPDSVDYIGLSRKTFLTALAAPQGEVVALAAVYPRFPO
RSEIYIDLAVSGEHRROGIALNALNLKHEMNALGAVIYIYQADYGDPPAVALTKL
GIREVMEHPDIDESTAT"
complement (4301..6394)
/gene="CD4-EGFP"
complement (4301..6394)
/note="CD4 and enhanced GFP fusion protein"
/codon_start=1
/transl_table=11
/product="CD4-EGFP"
/protein_id="AAR13802.1"
/db_xref="GI:38197737"
/translation="MNRGVPRHLLIVLQALLPAATQKKVVLGKKQDTVELCTAS
OKSIIQPHKNSNQIKIIGNQSFLTKGPSKLNDPADRSLSMDGNPFLITKNIKIE
DSDIYICEVDQKEVQLVREGLTNSDTHLQGSLLTULESPGSSPSVQCRPRG
KNIQGGKTLVSQLELODSGTWCTVLQNKQVEPKIDIVLAAPKASIVYKKEGEO
VEFSPLAFVTEKLTGSGELMWQBARASSKSMITFDLKNKESVYKRVYODPKOMGK
KLPLHLTLPOALPOYAGSGNLTLALEATGKXKHOENVLVNRAATOLOKNLTCBWGPT
SPKMLSLKLNKREKAVSRKRAVVLNPRAGMOCILSDSGOVLNLSNIRKLPWST
PVQPMALIVLGAVAGLLPFIGIGIFPCYCRHRRQARMSQIKLLSEKTKCQCPHR
FOKTCSPISVSKEBELFTGVVPIVELDGVNNGHPSVSGEBGATYKGLTKFTCT
TGKLPVWPVTLVTLTYGVQCFSRYPDMKQHPDFKSMPEGVQERTIFEDDNYK
TRAEVKFEGLDILVNRILEKGIIDFKEDGNIIGKLEYNLSHVYIMADOKNGIKYNF
KIRNIEDGVSOLADHYQONTPIGDVPILPNDNYLSTQSALSKDPENKRDHVLLEF
VTAADITLGMELVYK"
```

## ORIGIN

```
Query Match 43.4%; Score 624.4; DB 12; Length 7533;
Best Local Similarity 99.8%; Pred. No. 1.7e-134;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGTGCTGCAACTGGCGCTCTCCCA 60
Db 6394 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGTGCTGCAACTGGCGCTCTCCCA 6335
Qy 61 GCAGCACTCAGGGAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACGTACC 120
Db 6334 GCAGCACTCAGGGAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACGTACC 6275
Qy 121 TGTACAGCTTCCCAAGAGAGCATATTCACCTGGAATACTCCAAACAGATTAAG 180
Db 6274 TGTACAGCTTCCCAAGAGAGCATATTCACCTGGAATACTCCAAACAGATTAAG 6215
Qy 181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGATATGCGCT 240
Db 6214 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGATATGCGCT 6155
Qy 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTCCCTGATATCAAGAATCTTAAG 300
Db 6154 GACTCAAGAAAGCCTTTGGGACCAAGAACTCCCTGATATCAAGAATCTTAAG 6095
Qy 301 ATAGAAGACTCATGATCTTACATCTGTGAAGTGAGAGACCAAGAGAGAGTGCAATTG 360
Db 6094 ATAGAAGACTCATGATCTTACATCTGTGAAGTGAGAGACCAAGAGAGAGTGCAATTG 6035
Qy 361 CTAGTGTGGATTGATCTGCACTGTGACACCCACTGCTTACAGGGGCAAGACCTGACC 420
Db 6034 CTAGTGTGGATTGATCTGCACTGTGACACCCACTGCTTACAGGGGCAAGACCTGACC 5975
Qy 421 CTGACCTTGGAGAGCCCCCTGGTAGTACCCCTCAGTCAATGTAGAGTCCAGAGGGGT 480
Db 5974 CTGACCTTGGAGAGCCCCCTGGTAGTACCCCTCAGTCAATGTAGAGTCCAGAGGGGT 5915
Qy 481 AAAAACAATCAGGGGGGGAAGACCTCTCTCTGCTGCTGCTGAGTGTCCAGATAGTGGC 540
Db 5914 AAAAACAATCAGGGGGGGAAGACCTCTCTCTGCTGCTGCTGAGTGTCCAGATAGTGGC 5855
```

QY	541	ACCTGGAATGCACTGCTCTTTGCAAGAACCAAGAAAGTGAGAGTTCAAAATATGACATCGTG	600
Dd	5854	ACCTGGAATGCACTGCTCTTTGCAAGAACCAAGAAAGTGAGAGTTCAAAATATGACATCGTG	5795
QY	601	GTGCTAGCTTCCAGAAAGGCGCTCCGG	626
Dd	5794	GTGCTAGCTTCCAGAAAGGCGCTCCAG	5769
RESULT	11		
LOCUS	106223		
DEFINITION	106223	Sequence 1 from Patent EP 0313377.	1273 bp
ACCESSION	106223		linear
VERSION	106223.1	GI:590426	PAT 02-DEC-1994
KEYWORDS			
SOURCE		Unknown.	
ORGANISM		Unknown.	
REFERENCE		Unclassified.	
AUTHORS		1 (bases 1 to 1273)	
TITLE		Deen, K.C., Folema-Messerman, G.M., Inacker, R.H. and Sweet, R.W.	
JOURNAL		Process for purification of soluble T4	
FEATURES		Patent: EP 0313377-A1 1 26-Apr-1989;	
source		Location/Qualifiers	
		1..1273	

ORIGIN	/mol_type="unassigned DNA"									
Query Match	43.2%	Score 622.8;	DB 6;	Length 1273;						
Best Local Similarity	99.7%	Pred. No. 4.1e-134;								
Matches 624;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;						
QY	1	ATGAACCGGGGAGTCCCTTTTAGGACATCTTGCTCTGTGTGCTGCACATCGGCCTCCTCCCA	60							
DB	76	ATGAACCGGGGAGTCCCTTTTAGGACATCTTGCTCTGTGTGCTGCACATCGGCCTCCTCCCA	135							
QY	61	GCAGCCATCTCAGGGAAAGAAAGTGTGTGCTGGCCAAAAGGGGATATACGTGAACTGACC	120							
DB	136	GCAGCCATCTCAGGGAAACAAAGTGTGTGCTGGCCAAAAGGGGATATACGTGAACTGACC	195							
QY	121	TGTACAGCTTCCCAAGAAAGACATACATTTCCACTGAAAACTCCCAACAGATAAAG	180							
DB	196	TGTACAGCTTCCCAAGAAAGACATACATTTCCACTGAAAACTCCCAACAGATAAAG	255							
QY	181	ATTCTGGGAATCAGGGCTCCTCTTTAACTAAAGTCCATCCAGCTGAATGATCGCCT	240							
DB	256	ATTCTGGGAATCAGGGCTCCTCTTTAACTAAAGTCCATCCAGCTGAATGATCGCCT	315							
QY	241	GACTCAAGAAAGAGGCTTTGGGACCAAGAAAATTCCCCCTGATCATGAAGATCTTTAAG	300							
DB	316	GACTCAAGAAAGAGGCTTTGGGACCAAGAAAATTCCCCCTGATCATGAAGATCTTTAAG	375							
QY	301	ATAGAAGACTCAGATACTTACATCTGTGAAGTGAAGACCAAGAAAGAGGTGCATTTG	360							
DB	376	ATAGAAGACTCAGATACTTACATCTGTGAAGTGAAGACCAAGAAAGAGGTGCATTTG	435							
QY	351	CTAGTGTTCGATTACTACGCAACTCTGACACCCACTCTGTTCAGGGGCAAGCCTGACC	420							
DB	426	CTAGTGTTCGATTACTACGCAACTCTGACACCCACTCTGTTCAGGGGCAAGCCTGACC	495							
QY	421	CTGACCTTGGAGAGGCCCCCTGTGTATGAGCCCTCAGTGCATATGAAAGTCCAAAGGGGT	480							
DB	496	CTGACCTTGGAGAGGCCCCCTGTGTATGAGCCCTCAGTGCATATGAAAGTCCAAAGGGGT	555							
QY	481	AAAAACATACAGGGGGGGAAGACCCCTTCCGTGTCTCAGCTGGAAGCTCAGGATATGTGC	540							
DB	556	AAAAACATACAGGGGGGGAAGACCCCTTCCGTGTCTCAGCTGGAAGCTCAGGATATGTGC	615							
QY	541	ACCTGGACATGACATCTGTCTTGCAAGAACGAAGAAAGGTGAGATTAAAAATGACATCGTG	600							
DB	616	ACCTGGACATGACATCTGTCTTGCAAGAACGAAGAAAGGTGAGATTAAAAATGACATCGTG	675							

QY	601	GTGCTAGCTTCCAGAGGCTCCGG	626	
DB	676	GTGCTAGCTTCCAGAGGCTCCAG	701	
RESULT	12			
LOCUS	107147	1273 bp	DNA	linear
DEFINITION	Sequence 1 from Patent EP 0331356.			PAT 02-DEC-1994
ACCESSION	107147			
VERSION	107147.1	GI:590041		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	1 (bases 1 to 1273)			
AUTHORS	Archos, J., Clair, P. E., Fornwald, J. A., Brawner, M. E.,			
	Gorman, J. A., Sathie, G. W., Sweet, R. W. and Taylor, D. P.			
TITLE	Expression of HIV binding proteins			
JOURNAL	Patent: EP 0331356-A2 1 06-SEP-1989;			
FEATURES	Location/Qualifiers			
	1..1273			
source	/organism="unknown"			
	/mol_type="unassigned DNA"			
ORIGIN				

Query Match	Similarity	93.2%	Score 622.8	DB 6	Length 1273
Best Local	Similarity	99.7%	Pred. No. 4.1e-134		
Matches	Conservative	0	Mismatches	2	Indels
					Gaps
					0
QY	1	ATGAACCGGGAGTCCCTTTTAGGCACTGTGCTGTGCTGCAACTGGCGCTCTCCCA	60		
Db	76	ATGAACCGGGAGTCCCTTTTAGGCACTGTGCTGTGCTGCAACTGGCGCTCTCCCA	135		
QY	61	GCAGCACTCAGGGAAAGAAAGTGTCTGTGGCAAAAAGGGGATACAGTGAACCTACC	120		
Db	136	GCAGCACTCAGGGAAAGAAAGTGTGTGGCAAAAAGGGGATACAGTGAACCTACC	195		
QY	121	TGTACAGCTTCCAGAAAGAGAGATACAATTCACATGSAAAAACCTCAACAGATTAAG	180		
Db	196	TGTACAGCTTCCAGAAAGAGATACAAATTCACATGSAAAAACCTCAACAGATTAAG	255		
QY	181	ATTCTGGGAATATCAGGGCTCTCTTTTAACTTAAAGTTCATCAAGCTGAATGATCGGCT	240		
Db	256	ATTCTGGGAATATCAGGGCTCTCTTTTAACTTAAAGTTCATCAAGCTGAATGATCGGCT	315		
QY	241	GACTCAAGAAAGAGCTTTGGGACCAAGAAACTTCCCGGATCATCAAGATCTTTAAG	300		
Db	316	GACTCAAGAAAGAGCTTTGGGACCAAGAAACTTCCCGGATCATCAAGATCTTTAAG	375		
QY	301	ATAGAAGCTCAGATACTTACATCTGTGAAAGTGAAGACCAAGAAAGAGAGTGCATTG	360		
Db	376	ATAGAAGCTCAGATACTTACATCTGTGAAAGTGAAGACCAAGAAAGAGAGTGCATTG	435		
QY	361	CTAGGTTCCGATTTGACCTGCCAATCTGACACCAACCTGCTTCAAGGGCAGAGCTGACC	420		
Db	436	CTAGGTTCCGATTTGACCTGCCAATCTGACACCAACCTGCTTCAAGGGCAGAGCTGACC	495		
QY	421	CTGACCTTTGAGAGCCCTTGTTAGTAGCCCTTCAGTGCATATGAGAGTCCAAGGGT	480		
Db	496	CTGACCTTTGAGAGCCCTTGTTAGTAGCCCTTCAGTGCATATGAGAGTCCAAGGGT	555		
QY	481	AAAAACATACAGGGGGGGGAAAGACCTTCTCGTGTCTCAAGCTGAGAGCTTCAAGATAGTGGC	540		
Db	556	AAAAACATACAGGGGGGGGAAAGACCTTCTCGTGTCTCAAGCTGAGAGCTTCAAGATAGTGGC	615		
QY	541	ACCTGGAATATGACATGCTTTCAGAAACAGAAAGAGTGAATTCAAATATGACATCGTG	600		
Db	616	ACCTGGAATATGACATGCTTTCAGAAACAGAAAGAGTGAATTCAAATATGACATCGTG	675		
QY	601	GTGCTAGCTTTCCAGAAAGGCTCCGG	626		
Db	676	GTGCTAGCTTTCCAGAAAGGCTCCGG	701		

RESULT 13  
LOCUS AR067943 1304 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 28 from patent US 5851828.  
ACCESSION AR067943  
VERSION AR067943.1 GI:5999165  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1304)  
AUTHORS Seed,B., Banapour,B., Romeo,C. and Kolanus,W.  
TITLE Targeted cytolysis of HIV-infected cells by chimeric CD4  
JOURNAL receptor-bearing cells  
FEATURES  
LOCATION/Qualifiers  
1..1304  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 43.2%; Score 622.8; DB 6; Length 1304;  
Best Local Similarity 99.7%; Pred. No. 4.1e-134;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTGCTGCTGCAACTGCGCTCTCCCA 60  
DB 111 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTGCTGCTGCAACTGCGCTCTCCCA 170  
QY 61 GCAGCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC 120  
DB 171 GCAGCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC 230  
QY 121 TGTACAGCTTCCCAAGAAAGCATACATTCCTGGAATACTCCAACTGATAAG 180  
DB 231 TGTACAGCTTCCCAAGAAAGCATACATTCCTGGAATACTCCAACTGATAAG 290  
QY 181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGAATGCGCT 240  
DB 291 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGAATGCGCT 350  
QY 241 GACTCAAGAAAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 351 GACTCAAGAAAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 410  
QY 301 ATAGAAGCTCAGATTAATCACTGTGAAGTGAAGACCAAGAAAGGATGCAATTG 360  
DB 411 ATAGAAGCTCAGATTAATCACTGTGAAGTGAAGACCAAGAAAGGATGCAATTG 470  
QY 361 CTAGTGTTCGATTAAGTCACTGCACTGCACTGCTTCAAGGGGCAAGCTGACC 420  
DB 471 CTAGTGTTCGATTAAGTCACTGCACTGCACTGCTTCAAGGGGCAAGCTGACC 530  
QY 421 CTGACCTTGGAGAGGCCCCCTGTGTAGTACCCCTCAATGCAATGTAAGAGGCT 480  
DB 531 CTGACCTTGGAGAGGCCCCCTGTGTAGTACCCCTCAATGCAATGTAAGAGGCT 590  
QY 481 AAAAATATACAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCAGGATGTTG 540  
DB 591 AAAAATATACAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCAGGATGTTG 650  
QY 541 ACCTGACATGCACTGTCTTTCAGAAACCAAGAAAGGTGAGTTCAAAATAGACATG 600  
DB 651 ACCTGACATGCACTGTCTTTCAGAAACCAAGAAAGGTGAGTTCAAAATAGACATG 710  
QY 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626  
DB 711 GTGCTAGCTTTCAGAAAGGCTCCAG 736

RESULT 14

AR166802  
LOCUS AR166802 1304 bp DNA linear PAT 17-OCT-2001  
DEFINITION Sequence 28 from patent US 6284240.  
ACCESSION AR166802  
VERSION AR166802.1 GI:16243143  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1304)  
AUTHORS Seed,B., Banapour,B., Romeo,C. and Kolanus,W.  
TITLE Targeted cytolysis of HIV-infected cells by chimeric CD4  
JOURNAL receptor-bearing cells  
FEATURES  
LOCATION/Qualifiers  
1..1304  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 43.2%; Score 622.8; DB 6; Length 1304;  
Best Local Similarity 99.7%; Pred. No. 4.1e-134;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTGCTGCTGCAACTGCGCTCTCCCA 60  
DB 111 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTGCTGCTGCAACTGCGCTCTCCCA 170  
QY 61 GCAGCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC 120  
DB 171 GCAGCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC 230  
QY 121 TGTACAGCTTCCCAAGAAAGCATACATTCCTGGAATACTCCAACTGATAAG 180  
DB 231 TGTACAGCTTCCCAAGAAAGCATACATTCCTGGAATACTCCAACTGATAAG 290  
QY 181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGAATGCGCT 240  
DB 291 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGAATGCGCT 350  
QY 241 GACTCAAGAAAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 351 GACTCAAGAAAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 410  
QY 301 ATAGAAGCTCAGATTAATCACTGTGAAGTGAAGACCAAGAAAGGATGCAATTG 360  
DB 411 ATAGAAGCTCAGATTAATCACTGTGAAGTGAAGACCAAGAAAGGATGCAATTG 470  
QY 361 CTAGTGTTCGATTAAGTCACTGCACTGCACTGCTTCAAGGGGCAAGCTGACC 420  
DB 471 CTAGTGTTCGATTAAGTCACTGCACTGCACTGCTTCAAGGGGCAAGCTGACC 530  
QY 421 CTGACCTTGGAGAGGCCCCCTGTGTAGTACCCCTCAATGCAATGTAAGAGGCT 480  
DB 531 CTGACCTTGGAGAGGCCCCCTGTGTAGTACCCCTCAATGCAATGTAAGAGGCT 590  
QY 481 AAAAATATACAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCAGGATGTTG 540  
DB 591 AAAAATATACAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCAGGATGTTG 650  
QY 541 ACCTGACATGCACTGTCTTTCAGAAACCAAGAAAGGTGAGTTCAAAATAGACATG 600  
DB 651 ACCTGACATGCACTGTCTTTCAGAAACCAAGAAAGGTGAGTTCAAAATAGACATG 710  
QY 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626  
DB 711 GTGCTAGCTTTCAGAAAGGCTCCAG 736

RESULT 15  
AR062468  
LOCUS AR062468 1389 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 2 from patent US 5843728.

ACCESSION AR062468  
VERSION AR062468.1 GI:5990159  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 1389)  
AUTHORS Seed, B., Romeo, C. and Kolanus, W.  
TITLE Redirection of cellular immunity by receptor chimeras  
JOURNAL Patent: US 5843728-A 2 01-DEC-1998;  
FEATURES Location/Qualifiers  
1..1389  
source /organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 43.2%; Score 622.8; DB 6; Length 1389;  
Best Local Similarity 99.7%; Pred. No. 4,1e-134;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAACGGGGAGTCCCTTTTAGGCACTTGCTGGTGCACCTGGGCTCTCCCA 60  
DB 1 ATGAACGGGGAGTCCCTTTTAGGCACTTGCTGGTGCACCTGGGCTCTCCCA 60  
QY 61 GCAGCCACTCAGGGAAAGAAAGTGTCTGGGCAAAAAAGGGATACAGTGAACCTGACC 120  
DB 61 GCAGCCACTCAGGGAAAGAAAGTGTCTGGGCAAAAAAGGGATACAGTGAACCTGACC 120  
QY 121 TGTACAGCTTCCCGAAGAAAGAGCATCAATTCCACTGGAAAACTCCAACAGATTAAG 180  
DB 121 TGTACAGCTTCCCGAAGAAAGAGCATCAATTCCACTGGAAAACTCCAACAGATTAAG 180  
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGATCAATCAAGTGAATGAGGCT 240  
DB 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGATCAATCAAGTGAATGAGGCT 240  
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
QY 301 ATAGAACTCAGATTAATTAATTTGTAAGTGAAGACCAAGAAAGAGGTGCAATTG 360  
DB 301 ATAGAACTCAGATTAATTAATTTGTAAGTGAAGACCAAGAAAGAGGTGCAATTG 360  
QY 361 CTAGTGTTCGATTAATTAATTTGTAAGTGAAGACCAAGAGGTGCAAGCTGACC 420  
DB 361 CTAGTGTTCGATTAATTAATTTGTAAGTGAAGACCAAGAGGTGCAAGCTGACC 420  
QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGCATGTAGAGTCCAAGGGGT 480  
DB 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGCATGTAGAGTCCAAGGGGT 480  
QY 481 AAAAACAATACAGGGGGGAAAGACCTCTCCGTGTCACTGAGCTCCAGATATGGGC 540  
DB 481 AAAAACAATACAGGGGGGAAAGACCTCTCCGTGTCACTGAGCTCCAGATATGGGC 540  
QY 541 ACCTGGACATGCACTGTCTTGCAGAACCAAGAAAGGTGAGTCAAAATAGACATGCTG 600  
DB 541 ACCTGGACATGCACTGTCTTGCAGAACCAAGAAAGGTGAGTCAAAATAGACATGCTG 600  
QY 601 GTGCTAGCTTTCCAGAGGCTCCGG 626  
DB 601 GTGCTAGCTTTCCAGAGGCTCCAG 626

Search completed: November 22, 2004, 12:51:24  
Job time : 6284 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 14:00:14 ; Search time 754 Seconds  
(without alignments)  
10025.419 Million cell updates/sec

Title: US-09-936-702-4

Perfect score: 1440

Sequence: 1 atgacccggggagcctt.....gtcagcagtaataactg 1440

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*

1: geneeqn1980s:\*\n2: geneeqn1990s:\*\n3: geneeqn2000s:\*\n4: geneeqn2001as:\*\n5: geneeqn2001bs:\*\n6: geneeqn2002as:\*\n7: geneeqn2002bs:\*\n8: geneeqn2003as:\*\n9: geneeqn2003bs:\*\n10: geneeqn2003cs:\*\n11: geneeqn2003ds:\*\n12: geneeqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1440	100.0	1440	3	AAa54045 sCD4-SCFV
2	625	43.4	8911	12	Adn07735 Expressio
3	624.4	43.4	1377	4	AAf82582 Human CD4
4	624.4	43.4	1415	2	AAQ38761 sCD4-L2.1
5	624.4	43.4	1419	10	ADA44806 CD4/TCR C
6	624.4	43.4	1421	2	AAQ38759 sCD4-L1.1
7	624.4	43.4	1448	2	AAQ38759 sCD4-HAP
8	624.4	43.4	1714	3	AAa50662 DNA encod
9	624.4	43.4	1714	3	AAa44063 Human fus
10	624.4	43.4	1714	3	AAa48203 DNA seque
11	624.4	43.4	1742	3	AAa35205 Human ade
12	624.4	43.4	1742	3	AAf21327 Human low
13	624.4	43.4	1742	4	AAa165462 Nucleotid
14	624.4	43.4	1742	10	ABz97021 Human nuc
15	624.4	43.4	1742	11	AD131687 Human nuc
16	624.4	43.4	1742	11	ABD20870 Human pul
17	624.4	43.4	1910	3	AAa50663 DNA encod
18	624.4	43.4	1910	3	AAa44064 Human fus
19	624.4	43.4	1910	3	AAa48204 DNA seque
20	624.4	43.4	2150	3	AAa48205 DNA seque
21	624.4	43.4	2589	1	AAa90357 Genetic c

22	624.4	43.4	2589	3	AAa50661 DNA encod
23	624.4	43.4	2589	3	AAa44062 Human fus
24	624.4	43.4	2589	3	AAa48202 DNA seque
25	624.4	43.4	3133	1	AAa90356 Genetic c
26	624.4	43.4	3133	3	AAa50660 DNA encod
27	624.4	43.4	3133	3	AAa44061 Human fus
28	624.4	43.4	3133	3	AAa48201 DNA seque
29	624.4	43.4	6019	3	AAa35208 Human ade
30	624.4	43.4	6019	3	AAf21330 Human low
31	624.4	43.4	6019	10	ABz97024 Human nuc
32	624.4	43.4	6019	11	ABD20873 Human pul
33	622.8	43.2	1304	2	AAQ96103 CD4 doma
34	622.8	43.2	1304	2	AAf10797 CD4 D1-D4
35	622.8	43.2	1389	2	AAQ28705 DNA encod
36	622.8	43.2	1389	2	AAQ96123 T-cell re
37	622.8	43.2	1389	2	AAf10802 CD4:gamma
38	622.8	43.2	1389	2	AAf36759 CD4:FC re
39	622.8	43.2	1389	2	AAV70158 Chimeric
40	622.8	43.2	1599	2	AAQ28706 DNA encod
41	622.8	43.2	1599	2	AAf10803 CD4:eta f
42	622.8	43.2	1599	2	AAf36760 CD4:T-cel
43	622.8	43.2	1599	2	AAV70157 Chimeric
44	622.8	43.2	1727	2	AAQ96124 T-cell re
45	622.8	43.2	1728	2	AAQ28704 DNA encod

#### ALIGNMENTS

RESULT 1	AAA54045	standard: DNA; 1440 BP.
XX	AAA54045	
AC	AAA54045	
XX		
DT	08-FEB-2001	(first entry)
XX		
DE	sCD4-SCFV(17b)	HIV single chain antibody fusion protein.
XX		
KW	Fusion protein; HIV; human immunodeficiency virus; antibody; Fv; AIDS;	
KW	acquired immune deficiency syndrome; neutralisation; infection;	
KW	gene therapy; CD4; gp120; glycoprotein; resistance; vaccination;	
KW	binding domain; single chain antibody; chimera; chimeric protein; ds.	
XX		
OS	Human immunodeficiency virus.	
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..1440
FT		/*tag= a
FT		/product= "CD4-SCFV(17b) fusion protein."
FT		/note= "This sequence does not fully encode the cross
FT		referenced protein given in GENESQ record AAB00158. That
FT		protein sequence also comprises a 28 amino acid C-
FT		terminal peptide"
XX		
XX	WO200055207-A1.	
XX	21-SEP-2000.	
XX	16-MAR-2000; 2000WO-US006946.	
XX	16-MAR-1999; 99US-0124681P.	
XX	(USSH ) US NAT INST OF HEALTH.	
XX	Berger EA, Del Castillo CM;	
XX	WPI, 2000-638183/61.	
XX	P-PSDB; AAB00158.	
XX	Novel neutralizing bispecific fusion proteins effective in viral such as	
XX	HIV neutralization, comprises two different binding domains, inducing-	

PT binding domain and induced-binding domain functionally linked by linker.  
XX  
XX Claim 36; Page 47-48; 55pp; English.  
CC sCD4-sCFv(17b) is a neutralising bispecific fusion protein capable of  
CC binding to two sites of its target protein. The protein comprises a first  
CC binding domain capable of binding to an inducing site on the target  
CC protein, a second binding domain capable of forming neutralising complex  
CC with an induced epitope of the target protein and a linker connecting the  
CC binding domains. sCD4-sCFv(17b) comprises a soluble CD4 fragment  
CC (containing domains D1 and D2) fused to a single chain Fv portion of  
CC antibody 17b via a linker. sCD4-sCFv(17b), its variant, analogue or  
CC mimetic is used for inactivating gp120 protein of HIV, and for  
CC neutralising HIV. It is also used for blocking and preventing the binding  
CC of the viral or recombinant gp120 protein to soluble CD4 or lymphocyte  
CC CD4 and for inhibiting HIV replication. The chimeric proteins is  
CC therefore useful for treating HIV infection and also AIDS. It is are  
CC particularly useful in the prevention of infection during or immediately  
CC after HIV exposure (e.g., mother/infant transmission, post-exposure  
CC prophylaxis, and as a topical inhibitor) and for providing long term  
CC resistance to HIV infections and AIDS. Gene therapy is used to secrete  
CC the bispecific protein at mucosal surfaces, such as the vaginal, rectal  
CC or oral mucosa. The fusion proteins is highly potent, broadly cross-  
CC reactive with neutralising antibody with high in vivo activity and no Fc-  
CC mediated undesirable targeting properties. When the fusion protein is  
CC substantially derived from human proteins, it has minimal immunogenicity  
CC and toxicity in humans which is of great value in prevention of infection  
CC during or immediately after HIV exposure  
XX  
XX Sequence 1440 BP; 345 A; 354 C; 452 G; 289 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 1440; DB 3; Length 1440;  
Best Local Similarity 100.0%; Pred. No. 1.4e-311;  
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAACCGGAGAGTCCCTTTTAGGCACTTGTGTGTGTCGCAATGGCGCTCTCCCA 60  
DB 1 ATGAACCGGAGAGTCCCTTTTAGGCACTTGTGTGTGTCGCAATGGCGCTCTCCCA 60  
QY 61 GCAGCCCTCAGGGAGAAAGTGTGTGTGCGCAAAAAGGGGATACAGTGAACCTAACC 120  
DB 61 GCAGCCCTCAGGGAGAAAGTGTGTGTGCGCAAAAAGGGGATACAGTGAACCTAACC 120  
QY 121 TGTACAGCTTCCGAGAGAGAGCATCAATTCCACTGGAAGAACTCCACCAAGATAAG 180  
DB 121 TGTACAGCTTCCGAGAGAGAGCATCAATTCCACTGGAAGAACTCCACCAAGATAAG 180  
QY 181 ATTCTGGAAATCAGGCTCTCTTTAACTAAAGTCCATCAAGTGAATGATCGGCT 240  
DB 181 ATTCTGGAAATCAGGCTCTCTTTAACTAAAGTCCATCAAGTGAATGATCGGCT 240  
QY 241 GACTCAGAGAAAGCTTTGGGACCAAGAAACTCCCGTCGATCATCAAGAACTTAAG 300  
DB 241 GACTCAGAGAAAGCTTTGGGACCAAGAAACTCCCGTCGATCATCAAGAACTTAAG 300  
QY 301 ATGAGAACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGAGGTCAATTG 360  
DB 301 ATGAGAACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGAGGTCAATTG 360  
QY 361 CTAGTGTTCGATGACTGCCAATCTGACACCACTGCTTCAGGGGACAGACTGAAC 420  
DB 361 CTAGTGTTCGATGACTGCCAATCTGACACCACTGCTTCAGGGGACAGACTGAAC 420  
QY 421 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCTCAGTGCATATGAGAGTCCAAAGGAGT 480  
DB 421 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCTCAGTGCATATGAGAGTCCAAAGGAGT 480  
QY 481 AAAAACAATACAGGGGGAGAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGGC 540  
DB 481 AAAAACAATACAGGGGGAGAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGGC 540  
QY 541 ACCTGACATGCACTGTCTTGACAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 600  
DB 541 ACCTGACATGCACTGTCTTGACAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 600

DB 541 ACCTGACATGCACTGTCTTGACAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 600  
QY 601 GTGTAGACTTCCAGAAAGGCTCCGGAGGTGGCGATAGTGGGGAGGCGTTGACGCGA 660  
DB 601 GTGTAGACTTCCAGAAAGGCTCCGGAGGTGGCGATAGTGGGGAGGCGTTGACGCGA 660  
QY 661 GGTGATCCGGTGGCGAGGGTGGGGCGGGGTGGAAGCGGGGTGGCGCTCCGAGGC 720  
DB 661 GGTGATCCGGTGGCGAGGGTGGGGCGGGGTGGAAGCGGGGTGGCGCTCCGAGGC 720  
QY 721 GAGGTTCAAGTGCAGTGTCTCGATCTGGGGCTGAGGTGAAGAAAGCTGGGTCTCG 780  
DB 721 GAGGTTCAAGTGCAGTGTCTCGATCTGGGGCTGAGGTGAAGAAAGCTGGGTCTCG 780  
QY 781 GTAAGGTCTCTGCAAGGCTCTGAGAGACCTTATCATGATATGTTTACTGGGTG 840  
DB 781 GTAAGGTCTCTGCAAGGCTCTGAGAGACCTTATCATGATATGTTTACTGGGTG 840  
QY 841 CGAAGGCTCTGAGACAGGCTTTGAGTGAAGTGAAGATCATCACTATCTTGATGA 900  
DB 841 CGAAGGCTCTGAGACAGGCTTTGAGTGAAGTGAAGATCATCACTATCTTGATGA 900  
QY 901 GCACACTACGCAACCGCACTCCAGGGCAGAGTACAGATTACCGGACAAAGTCCAGAGC 960  
DB 901 GCACACTACGCAACCGCACTCCAGGGCAGAGTACAGATTACCGGACAAAGTCCAGAGC 960  
QY 961 ACAGTCTACCTGAGCTGCGGAATCTTAAGATCTGACGATCGGCGGATATTTCTGTGCG 1020  
DB 961 ACAGTCTACCTGAGCTGCGGAATCTTAAGATCTGACGATCGGCGGATATTTCTGTGCG 1020  
QY 1021 GAGGTGACAGGAGAGAGCGGAGCGAGCGGGGAATATGATTAATATGAGTTTCTGAAACAT 1080  
DB 1021 GAGGTGACAGGAGAGAGCGGAGCGGAGCGGGGAATATGATTAATATGAGTTTCTGAAACAT 1080  
QY 1081 TGGGGCCAGGAAACCTGTGTCAAGTCACTCAGTGTGGGTGGTCCGAGGTGTGGG 1140  
DB 1081 TGGGGCCAGGAAACCTGTGTCAAGTCACTCAGTGTGGGTGGTCCGAGGTGTGGG 1140  
QY 1141 AGCGGTGGCGGAGATCTGAATCTGAGTTGACCAAGTCTCAGCCACCTGTCTGTGTCT 1200  
DB 1141 AGCGGTGGCGGAGATCTGAATCTGAGTTGACCAAGTCTCAGCCACCTGTGTGTCT 1200  
QY 1201 CCAAGGAAAGACCACTCTCTGCAAGGCGCAGTGAAGTGTAGTGAACCTTAAGCC 1260  
DB 1201 CCAAGGAAAGACCACTCTCTGCAAGGCGCAGTGAAGTGTAGTGAACCTTAAGCC 1260  
QY 1261 TGGTACCAAGAAACCTGTGCAAGGCTCCAGGCTCTCTATATATGTGTATCCACAGG 1320  
DB 1261 TGGTACCAAGAAACCTGTGCAAGGCTCCAGGCTCTCTATATATGTGTATCCACAGG 1320  
QY 1321 GCCACCGGTGTCCAGGCAAGTTCAGTGCAGTGGGTCTGGGGCAGAAATCACTTCACC 1380  
DB 1321 GCCACCGGTGTCCAGGCAAGTTCAGTGCAGTGGGTCTGGGGCAGAAATCACTTCACC 1380  
QY 1381 ATCAGACGCTGCAAGTCTGAAGATTTTGCAGTTTATTAATCTGACAGCATCAATACTGG 1440  
DB 1381 ATCAGACGCTGCAAGTCTGAAGATTTTGCAGTTTATTAATCTGACAGCATCAATACTGG 1440  
RESULT 2  
ADN07735  
ID ADN07735 standard; DNA; 8911 BP.  
XX  
XX ADN07735;  
AC  
XX 15-JUL-2004 (first entry)  
XX  
XX Expression vector DNA.  
XX  
XX Immunogenic complex; gp120; CD4; HIV; HIV infection; anti-HIV; circular;  
KM cyclic; ds.  
XX  
XX Synthetic.

Query Match	43.4%	Score 625	DB 12	Length 8911
Best Local Similarity	100.0%	Pred. No. 1.7e-129		
Matches 625; Conservative	0	Mismatches 0	Indels 0	Gaps 0
Sequence 8911 BP; 2188 A; 2221 C; 2227 G; 2275 T; 0 U; 0 Other;				
1 ATGAACCGGGAGTCCCTTTTAGGACATTCCTTGATGCTCAACTGGCGCTCCCA	60			
3815 ATGAACCGGGAGTCCCTTTTAGGACATTCCTTGATGCTCAACTGGCGCTCCCA	3877			
61 GCAGCACTCAGGAGAAAGATGTGTCTGGGCAAAAAAGGGATATCACTGGAAGTACC	120			
3875 GCAGCACTCAGGAGAAAGATGTGTCTGGGCAAAAAAGGGATATCACTGGAAGTACC	393			
121 TGTAAAGCTTCCCAAGAAAGATCAATTCCTCACTGGAAGAACTCCCAACAGATAAG	180			
3935 TGTAAAGCTTCCCAAGAAAGATCAATTCCTCACTGGAAGAACTCCCAACAGATAAG	399			
181 ATTCTGGGAAATCAGGGCTCTCTTTAACTAAAGGTCATCCAGGTGAATGTCGGCT	240			
3995 ATTCTGGGAAATCAGGGCTCTCTTTAACTAAAGGTCATCCAGGTGAATGTCGGCT	405			
241 GACTCAAGAAAGAGCTTTTGGAGCAAGAGAACTTCCCGCTGATCATCAAGATCTTAA	300			
4055 GACTCAAGAAAGAGCTTTTGGAGCAAGAGAACTTCCCGCTGATCATCAAGATCTTAA	411			
301 ATAGAAGACTCGAATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG	360			

Dd		4115	ATGAGAAGCTCAGATTACTTACATCTGTGAAGTGAGAGCAACAAGAGAGGTGCATTGTG	4114
Oy		361	CTAGTGTTCGGATTGACTGCCAAGCTGTGACACCACCTGCTTCAGGGCGAGAGCTGACC	420
Dd		4175	CTTAGTTGTCGGATTGACTGCCAACTGTGACACCACCTGCTTCAGGGCGAGAGCTGACC	4234
Oy		421	CTGACCTTTGGAGAACCCCCCTGGTAAGTAAAGCCCTCATGTGAAATGTAGAGTCCAAAGGGGT	480
Dd		4235	CTGACCTTTGGAGAACCCCCCTGGTAAGTAAAGCCCTCATGTGAAATGTAGAGTCCAAAGGGGT	4294
Oy		481	AAAAACATACAGGGGGGGGAAAGACCCCTGCTCGGTGTCTAGCTGAGCTCCAGATATGAGGC	540
Dd		4295	AAAAACATACAGGGGGGGGAAAGACCCCTGCTCGGTGTCTAGCTGAGCTCCAGATATGAGGC	4354
Oy		541	ACCTGGACATGACATCTGCTTGTGCAGAACCAAGAAGGTGAGTTCAAATAATGACATCTGTG	600
Dd		4355	ACCTGGACATGACATCTGCTTGTGCAGAACCAAGAAGGTGAGTTCAAATAATGACATCTGTG	4414
Oy		601	GTCCTACCTTTCAGGAAGCCTCCG	625
Dd		4415	GTCCTACCTTTCAGGAAGCCTCCG	4439
<b>RESULT 3</b>				
ID	AAF82582		standard; cdNA; 1377 BP.	
XX	AAF82582;			
DT	18-JUN-2001	(first entry)		
DE	Human CD4 gene T4.			
KW	Human; CD4; T4; CD4 fusion protein; oligomerisation;			
KM	receptor-ligand interaction inhibition; surface plasmon resonance; SPR;			
KV	T cell receptor binding; MHC binding; carcinoma; autoimmune disease;			
KX	muciple sclerosis; human immunodeficiency virus; HIV; diabetes;			
RW	rheumatoid arthritis; immune disorder; ss.			
OS	Homo sapiens.			
XX				
FH	Key	Location/Qualifiers		
FT	CDS	1..1377		
FT		/*tag= a		
FT	sig_peptide	/product= "human CD4"		
FT		1..75		
FT	mat_peptide	/*tag= b		
FT		76..1374		
FT		/*tag= c		
XX				
PN	MO200122084-A2.			
XX				
PD	29-MAR-2001.			
XX				
PF	18-SEP-2000; 2000OWO-GB003579.			
XX				
PR	21-SEP-1999; 99GB-00022352.			
XX				
PA	(AVID-) AVIDEX LTD.			
P1	Jakobsen BK;			
XX				
DR	WP1; 2001-273470/28.			
XX	P-PSDB; AAB81502.			
PT	Sequential screening of candidate compounds library for those which			
FT	inhibit binding of low affinity receptor-ligand interaction having fast			
XX	binding kinetics, using interfacial optical assay.			
PS	Disclosure; Fig 13; 91pp; English.			
XX	The present sequence encodes human CD4. Human CD4 extracellular domains 1			
CC	and 2 were used in the construction of CD4 oligomerisation fusion			







```

RESULT 6
AAQ38760 ID AAQ38760 standard; DNA; 1421 BP.
XX
XX AAQ38760;
AC
XX 25-MAR-2003 (revised)
DT 22-JUL-1993 (first entry)
XX
XX sCD4-L1 lysosomal targeting fusion gene.
DE
XX
XX Soluble CD4; HIV; Human immunodeficiency virus; envelope; glycoprotein;
KM polymerase chain reaction; lysosomal membrane protein; lamp-1; ss.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH misc_feature 1..6
FT /*tag= a
FT /note= "EcoRI restriction site"
FT misc_feature 1275..1280
FT /*tag= b
FT /note= "XbaI restriction site"
FT misc_feature 1416..1421
FT /*tag= c
FT /note= "Sali restriction site"
XX
XX WO9306216-A1.
XX
XX 01-APR-1993.
XX
XX 22-SEP-1992; 92MO-US008090.
XX
XX 26-SEP-1991; 91US-00766963.
XX
XX (OKLA-) OKLAHOMA MED RES FOUND.
XX
XX Tang JUN, Lin XL;
XX
XX WPI; 1993-117537/14.
XX
XX New fusion protein used in gene therapy for treating AIDS - comprises
PT protein which binds to retroviral envelope protein which targets fusion
PT protein to lysosome.
XX
XX Claim 12; Page 13-14; 47pp; English.
XX
XX Primers P-1 and P-2 (AAQ38748 and AAQ38749) were used to amplify an EcoRI
CC -XbaI fragment of plasmid pT48 containing the sCD4 sequence. The PCR
CC product was ligated to an XbaI-Sali fragment containing the sequence
CC encoding lysosomal membrane protein (LAMP-1) (PCR amplified from a human
CC liver lambda gpi0 library using primers AAQ38754 and AAQ38755). The
CC resulting fusion sequence sCD4-L1 can be used to interfere with the
CC normal function of HIV and to direct the virus to lysosomes. (Updated on
CC 25-MAR-2003 to correct PN field.)
XX
XX
XX Sequence 1421 BP; 355 A; 382 C; 401 G; 283 T; 0 U; 0 Other;
SQ
Query Match 43.4%; Score 624.4; DB 2; Length 1421;
Best Local Similarity 99.8%; Pred. No. 1.6e-129;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAACGGGGAGGCCCTTTAGGCACTTGCTTGAGGCTGCAAACTGGGCGCTCCTCCA 60
DB 81 ATGAACGGGGAGGCCCTTTAGGCACTTGCTTGAGGCTGCAAACTGGGCGCTCCTCCA 140
QY 61 GCAAGCACTCAGGAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120
DB 141 GCAAGCACTCAGGAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 200
QY 121 TGTACAGCTTCCGAGAAAGAGATACATTCCTCAGTGAAGAACTCCACAGATTAAG 180
DB 201 TGTACAGCTTCCGAGAAAGAGATACATTCCTCAGTGAAGAACTCCACAGATTAAG 260

```

```

QY 181 ATTCTGGAAATCAGGGCTCTCTTTAACTAAAGTCCATCCAAAGCTGATATCGGCT 240
DB 261 ATTCTGGAAATCAGGGCTCTCTTTAACTAAAGTCCATCCAAAGCTGATATCGGCT 320
QY 241 GACTCAAGAAAGCCTTTGGACCAAGAAATCTCCCTGATCATCAAGATCTTTAAG 300
DB 321 GACTCAAGAAAGCCTTTGGACCAAGAAATCTCCCTGATCATCAAGATCTTTAAG 380
QY 301 ATGAAGACTCAGATTCCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 360
DB 381 ATGAAGACTCAGATTCCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 440
QY 361 CTAGTGTTCGATTTGACTGCAACTGTGACACCCACTGCTTCAAGGGGACCTGACC 420
DB 441 CTAGTGTTCGATTTGACTGCAACTGTGACACCCACTGCTTCAAGGGGACCTGACC 500
QY 421 CTGACCTTGGAGAGCCCCCTGGTAGTAGGCCCTCAGTGCATATGTAGAGTCCAAAGGGGT 480
DB 501 CTGACCTTGGAGAGCCCCCTGGTAGTAGGCCCTCAGTGCATATGTAGAGTCCAAAGGGGT 560
QY 481 AAAAACAATCAGGGGGGGAAGACCTCTCCGTCTCAGCTGAGCTCCAGATAGTGGC 540
DB 561 AAAAACAATCAGGGGGGGAAGACCTCTCCGTCTCAGCTGAGCTCCAGATAGTGGC 620
QY 541 ACCTGACATGACATGCTCTTGCAGAAACCAAGAAAGTGAAGTTCAAAATAGACATCGTG 600
DB 621 ACCTGACATGACATGCTCTTGCAGAAACCAAGAAAGTGAAGTTCAAAATAGACATCGTG 680
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
DB 681 GTGCTAGCTTTCCAGAAAGGCTCCAG 706

RESULT 7
AAQ38759 ID AAQ38759 standard; DNA; 1448 BP.
XX
XX AAQ38759;
AC
XX 25-MAR-2003 (revised)
DT 22-JUL-1993 (first entry)
XX
XX sCD4-HAP lysosomal targeting fusion gene.
DE
XX
XX Soluble CD4; HIV; Human immunodeficiency virus; envelope; ss;
KM glycoprotein; polymerase chain reaction; human acid phosphatase.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH misc_feature 1..6
FT /*tag= a
FT /note= "EcoRI restriction site"
FT misc_feature 1444..1448
FT /*tag= b
FT /note= "Sali restriction site"
XX
XX WO9306216-A1.
XX
XX 01-APR-1993.
XX
XX 22-SEP-1992; 92MO-US008090.
XX
XX 26-SEP-1991; 91US-00766963.
XX
XX (OKLA-) OKLAHOMA MED RES FOUND.
XX
XX Tang JUN, Lin XL;
XX
XX WPI; 1993-117537/14.
XX
XX New fusion protein used in gene therapy for treating AIDS - comprises
PT

```

PT protein which binds to retroviral envelope protein which targets fusion  
 PT protein to lysosome.

PS Claim 12, Page 12-13; 47pp; English.

XX Primers P-1 and P-2 (AAQ38748 and AAQ38749) were used to amplify an EcoRI  
 CC -XbaI fragment of plasmid pT48 containing the scd4 sequence. The PCR  
 CC product was ligated to an XbaI-SalI fragment containing the sequence  
 CC encoding human acid phosphatase (HAP) lysosomal targeting protein (PCR  
 CC amplified from a human liver lambda gp10 library using primers AAQ38752  
 CC and AAQ38753). The resulting fusion sequence scd4-HAP can be used to  
 CC interfere with the normal function of HIV and to direct the virus to  
 CC lysosomes. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 1448 BP; 357 A; 395 C; 404 G; 292 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 2; Length 1448;

Best Local Similarity 99.8%; Pred. No. 1.6e-129;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGTCGCAACTGGCGCTCTCCCA 60  
 DB 81 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGTCGCAACTGGCGCTCTCCCA 140  
 QY 61 GGAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACGTGAACTGACC 120  
 DB 141 GGAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACGTGAACTGACC 200  
 QY 121 TGTACAGCTTCCAGAAAGAGCATATTCACACTGGAAAACTCCACAGATTAAG 180  
 DB 201 TGTACAGCTTCCAGAAAGAGCATATTCACACTGGAAAACTCCACAGATTAAG 260  
 QY 181 ATTCTGGGAATCAGGGCTCTTCTTAAGTAAAGTCCATCCAGCTGAATATGCGCT 240  
 DB 261 ATTCTGGGAATCAGGGCTCTTCTTAAGTAAAGTCCATCCAGCTGAATATGCGCT 320  
 QY 241 GACTCAAGAAAGCTTTGGGACCAAGAAACTCCCTGATATCAAGAACTTTAAG 300  
 DB 321 GACTCAAGAAAGCTTTGGGACCAAGAAACTCCCTGATATCAAGAACTTTAAG 380  
 QY 301 ATAGAAGACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGGTGCAATTG 360  
 DB 381 ATAGAAGACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGGTGCAATTG 440  
 QY 361 CTAGTGTTCGATTAATCACTGCAACCTGCTTCAAGGGGAGAGCTTGACC 420  
 DB 441 CTAGTGTTCGATTAATCACTGCAACCTGCTTCAAGGGGAGAGCTTGACC 500  
 QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATATGATGAGGGT 480  
 DB 501 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATATGATGAGGGT 560  
 QY 481 AAAAATCATTCAGAGGGGGAAGACCTCTGCTGTCTCAGCTGAGCTCCAGATATGAGC 540  
 DB 561 AAAAATCATTCAGAGGGGGAAGACCTCTGCTGTCTCAGCTGAGCTCCAGATATGAGC 620  
 QY 541 ACCTGAGACATGACATGTCTTTCAGAAACCAAGAAAGGTGAGTTCAAAATAGACATCTGTG 600  
 DB 621 ACCTGAGACATGACATGTCTTTCAGAAACCAAGAAAGGTGAGTTCAAAATAGACATCTGTG 680  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
 DB 681 GTGCTAGCTTTCCAGAAAGGCTCCAG 706

RESULT 8  
 AAA50662  
 ID AAA50662 standard; DNA; 1714 BP.

AC AAA50662;

XX 09-JAN-2001 (first entry)

DT

XX

DE DNA encoding CD4-IgM fusion protein CH4Mmu.

XX CD4; IgM; human; CD4Mmu; fusion protein; immunoglobulin; HIV; SIV; gp120;  
 KW therapy; diagnosis; ds.

XX Homo sapiens.

OS Location/Qualifiers

FT exon

FT Key

FT CDS

FT intron

FT exon

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGTCGCAACTGGCGCTCTCCCA 60  
 DB 111 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGTCGCAACTGGCGCTCTCCCA 170  
 QY 61 GGAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACGTGAACTGACC 120

Query Match 43.4%; Score 624.4; DB 3; Length 1714;  
 Best Local Similarity 99.8%; Pred. No. 1.6e-129;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Example 1; Col 41-50; 39pp; English.

The present sequence is that of DNA encoding fusion protein CD4Mmu (see  
 CC AAB19509) comprising the extracellular portion of CD4, which binds to HIV  
 CC gp120, linked at its C-terminus to a human IgM heavy chain polypeptide.  
 CC To obtain the construct, DNA encoding CD4 was linked to IgM DNA at the  
 CC Mst2 site upstream of the CH1 region. A plasmid containing this genetic  
 CC construct is deposited in Escherichia coli MC1061/P3 as ATCC 67609.

CC Fusion protein CD4Mmu and a nucleic acid encoding it are claimed. Also  
 CC claimed are a vector comprising the nucleic acid, and a method of  
 CC producing the fusion protein in secreted form using a transformed host  
 CC cell. The fusion protein may further comprise a therapeutic agent,  
 CC radiolabel or NMR imaging agent. The fusion protein can be administered  
 CC to an animal (including humans) for treatment of HIV or SIV infection.  
 CC and can also be used in assays for HIV or SIV, imaging and tissue stains.  
 CC IgM fusion proteins such as CD4Mmu provide complement-mediated immunity

XX Sequence 1714 BP; 411 A; 503 C; 467 G; 333 T; 0 U; 0 Other;

```
Db 171 GCAGCCACTCAGGAAAGAAAGTGTCTGTGGCAAAAAAGGGATACAGTGAATGACC 230
Qy 121 TGTACAGCTTCCCAAGAAAGAGATACATTCACCTGGAAGAACTCCACCATTAAG 180
Db 231 TGTACAGCTTCCCAAGAAAGAGATACATTCACCTGGAAGAACTCCACCATTAAG 290
Qy 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGTGAATGATCGGCT 240
Db 291 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGTGAATGATCGGCT 350
Qy 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
Db 351 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 410
Qy 301 ATGAAAGACTCAATTAATCTGTGAAAGTGAAGACCAAGAGAGAGTGCATTTG 360
Db 411 ATGAAAGACTCAATTAATCTGTGAAAGTGAAGACCAAGAGAGAGTGCATTTG 470
Qy 361 CTAGTGTTCGGAATGACTGCAACTGACCACTGCTTCAGGGGACAGGCTGACC 420
Db 471 CTAGTGTTCGGAATGACTGCAACTGACCACTGCTTCAGGGGACAGGCTGACC 530
Qy 421 CTGACCTTGAGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGGT 480
Db 531 CTGACCTTGAGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGGT 590
Qy 481 AAAAACAATACAGGGGGGGAAGACCTCTCTGTCTCACTGAGCTCCAGATATGTGGC 540
Db 591 AAAAACAATACAGGGGGGGAAGACCTCTCTGTCTCACTGAGCTCCAGATATGTGGC 650
Qy 541 ACCTGACATGCACTGTCTTGCAAGACCAAGAAAGGTGAGTCAAAATAGATGCTG 600
Db 651 ACCTGACATGCACTGTCTTGCAAGACCAAGAAAGGTGAGTCAAAATAGATGCTG 710
Qy 601 GTGCTAGCTTTCCAGAAAGGCTTCGG 626
Db 711 GTGCTAGCTTTCCAGAAAGGCTTCAG 736

RESULT 9
AAZ4063
ID AAZ44063 standard; DNA; 1714 BP.
XX
AC AAZ44063;
XX
DT 23-MAR-2000 (first entry)
XX
Human fusion protein CD4Mg encoding DNA.
XX
KW Fusion protein; human; CD4; IGM; immunoglobulin; gp120;
KM anti-human immunodeficiency virus; CD4Mg; ds.
XX
OS Homo sapiens.
XX
XX Synthetic.
XX
PN US6004781-A.
XX
PD 21-DEC-1999.
XX
PF 04-FEB-1994; 94US-00191708.
XX
PR 22-JAN-1988; 88US-00147351.
PR 23-JAN-1989; 89US-00299596.
PR 09-JUN-1992; 92US-00896781.
PR 12-APR-1993; 93US-00057952.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Seed B;
XX
WPI; 2000-085792/07.
XX
DR P-PSDB; AAY51080.
```

```
XX
PT Fusion protein useful for the treatment of human immunodeficiency virus.
XX
PS Example 1; Col 41-50; 39pp; English.
XX
CC This invention describes a novel nucleic acid (I) encoding a fusion
CC protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)
CC and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light
CC chain (III). The products of the invention have anti-human
CC immunodeficiency virus (HIV) activity and are capable of binding to
CC gp120. The fusion protein is useful for treating human immunodeficiency
CC virus (HIV) or similar immunodeficiency virus (SIV). This sequence encodes
CC the fusion protein CD4Mg which is constructed from CD4 linked to human
CC IGM upstream of the CH1 region
XX
SQ Sequence 1714 BP; 411 A; 503 C; 467 G; 333 T; 0 U; 0 Other;
Query Match 43.4%; Score 624.4; DB 3; Length 1714;
Best Local Similarity 99.8%; Pred. No. 1,6e-129;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGAACCGGGAGATCCCTTTTAAAGCACTTGCTGTGCAACTGGGGCTCTCCCA 60
Db 111 ATGAACCGGGAGATCCCTTTTAAAGCACTTGCTGTGCAACTGGGGCTCTCCCA 170
Qy 61 GCAGCCACTCAGGAAAGAAAGTGTGCTGGCAAAAAAGGGATACAGTGAATGACC 120
Db 171 GCAGCCACTCAGGAAAGAAAGTGTGCTGGCAAAAAAGGGATACAGTGAATGACC 230
Qy 121 TGTACAGCTTCCCAAGAAAGAGATACATTCACCTGGAAGAACTCCACCATTAAG 180
Db 231 TGTACAGCTTCCCAAGAAAGAGATACATTCACCTGGAAGAACTCCACCATTAAG 290
Qy 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGTGAATGATCGGCT 240
Db 291 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGTGAATGATCGGCT 350
Qy 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
Db 351 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 410
Qy 301 ATGAAAGACTCAATTAATCTGTGAAAGTGAAGACCAAGAGAGAGTGCATTTG 360
Db 411 ATGAAAGACTCAATTAATCTGTGAAAGTGAAGACCAAGAGAGAGTGCATTTG 470
Qy 361 CTAGTGTTCGGAATGACTGCAACTGACCACTGCTTCAGGGGACAGGCTGACC 420
Db 471 CTAGTGTTCGGAATGACTGCAACTGACCACTGCTTCAGGGGACAGGCTGACC 530
Qy 421 CTGACCTTGAGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGGT 480
Db 531 CTGACCTTGAGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGGT 590
Qy 481 AAAAACAATACAGGGGGGGAAGACCTCTCTGTCTCACTGAGCTCCAGATATGTGGC 540
Db 591 AAAAACAATACAGGGGGGGAAGACCTCTCTGTCTCACTGAGCTCCAGATATGTGGC 650
Qy 541 ACCTGACATGCACTGTCTTGCAAGACCAAGAAAGGTGAGTCAAAATAGATGCTG 600
Db 651 ACCTGACATGCACTGTCTTGCAAGACCAAGAAAGGTGAGTCAAAATAGATGCTG 710
Qy 601 GTGCTAGCTTTCCAGAAAGGCTTCGG 626
Db 711 GTGCTAGCTTTCCAGAAAGGCTTCAG 736

RESULT 10
AAZ48203
ID AAZ48203 standard; DNA; 1714 BP.
XX
AC AAZ48203;
XX
DT 14-MAR-2000 (first entry)
```

XX DNA sequence encoding CD4-Ig fusion protein CD4Mmu.  
 DE HIV, extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;  
 KM secreted protein; HIV infection; medicament; ds.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN CA1340741-C.  
 XX  
 PD 14-SEP-1999.  
 XX  
 PF 20-JAN-1989; 89CA-00588749.  
 XX  
 PR 20-JAN-1989; 89CA-00588749.  
 XX  
 PA (GENE) GEN HOSPITAL CORP.  
 XX  
 PI Seed B;  
 XX  
 DR WPI, 2000-063015/06.  
 XX  
 PT P-PSDB; AAY59170.  
 XX  
 PT New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in  
 the treatment of HIV or simian immunodeficiency virus infections.  
 XX  
 PS Example 1; Page 47-53; 89pp; English.  
 XX  
 CC The invention provides a fusion gene encoding a fusion protein that  
 comprises an extracellular CD4 DNA sequence or its fragment which binds  
 to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA  
 CC sequence of an Ig heavy or light chain, where the DNA sequence encoding  
 the variable region has been replaced with the DNA sequence which encodes  
 CC extracellular CD4 or its gp120 binding fragment. The fusion protein is  
 CC capable of being secreted. The fusion proteins are useful for treating  
 CC HIV or SIV infections in animals, preferably humans. They are also useful  
 CC for producing medicaments which can be used for treating HIV or SIV  
 CC infections in humans. The present sequence represents the DNA which  
 CC encodes the fusion protein CD4Mmu where the CD4 is linked to human IgG1  
 CC at the Mst2 site upstream of the CH1 region  
 CC  
 SQ Sequence 1714 BP; 411 A; 503 C; 467 G; 333 T; 0 U; 0 Other;  
 Query Match 43.4%; Score 624.4; DB 3; Length 1714;  
 Best Local Similarity 99.8%; Pred. No. 1.6e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTCGTCGCACTGGCGCTCCGCCA 60  
 DB 111 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTCGTCGCACTGGCGCTCCGCCA 170  
 QY 61 GCAGCCACTCAGGGAAGAAAGTGTGTGCTGGCAAAAAAGGATATCAGTGAAGTGAAC 120  
 DB 171 GCAGCCACTCAGGGAAGAAAGTGTGTGCTGGCAAAAAAGGATATCAGTGAAGTGAAC 230  
 QY 121 TGTATAGCTTCCCAAGAAAGCATATCAATTCCTGGAATAAATCCCAACCAATTAAG 180  
 DB 221 TGTATAGCTTCCCAAGAAAGCATATCAATTCCTGGAATAAATCCCAACCAATTAAG 290  
 QY 181 ATTCGGGAATCAGGGCTCCTTTTAACTAAGTCCATCCAACTGAATATATGGCCT 240  
 DB 291 ATTCGGGAATCAGGGCTCCTTTTAACTAAGTCCATCCAACTGAATATATGGCCT 350  
 QY 241 GACTCAAGAAAGAGCTTTGGAGCAAGAAATTCCTCCCTGATCATCAAGATCTTAAG 300  
 DB 351 GACTCAAGAAAGAGCTTTGGAGCAAGAAATTCCTCCCTGATCATCAAGATCTTAAG 410  
 QY 301 ATAGAAGCTCAGATTAATCACTCTGTGAAGTGAAGACCAAGAGAGAGAGTGAATTTG 360  
 DB 411 ATAGAAGCTCAGATTAATCACTCTGTGAAGTGAAGACCAAGAGAGAGTGAATTTG 470  
 QY 361 CTAGTGTTCGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 420

DB 471 CTAGTGTTCGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 530  
 QY 421 CTAGTGTTCGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 480  
 DB 531 CTAGTGTTCGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 590  
 QY 481 AAAAATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
 DB 591 AAAAATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 650  
 QY 541 ACCTGACATGACATGCTCTTTCGAGAACCAAGAAAGTGAAGTGAAGTGAAGTGAAG 600  
 DB 651 ACCTGACATGACATGCTCTTTCGAGAACCAAGAAAGTGAAGTGAAGTGAAGTGAAG 710  
 QY 601 GTGCTAGCTTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 626  
 DB 711 GTGCTAGCTTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 736  
 RESULT 11  
 AAA35205  
 ID AAA35205 standard; DNA; 1742 BP.  
 XX  
 AC AAA35205;  
 XX  
 DT 28-JUL-2000 (first entry)  
 XX  
 DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:79.  
 XX  
 KW Human; adenosine receptor; low adenosine antisense oligonucleotide;  
 KW phosphorothioate; impaired respiration; inflammation; allergy;  
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;  
 KW antiallergic; anticholinergic; cytotatic; analgesic; impaired airway;  
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;  
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;  
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;  
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; se.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200009525-A2.  
 XX  
 PD 24-FEB-2000.  
 XX  
 PF 03-AUG-1999; 99WO-US017712.  
 XX  
 PR 03-AUG-1999; 98US-0095212P.  
 XX  
 PA (UYEC-) UNIV EAST CAROLINA.  
 XX  
 PI Nyce JW;  
 XX  
 DR WPI, 2000-205971/18.  
 XX  
 PT New antisense oligonucleotides useful for treating e.g. pulmonary  
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,  
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or  
 PT cancers.  
 XX  
 PS Disclosure; Page 1245; 1343pp; English.  
 XX  
 CC The present invention describes a new composition comprising an antisense  
 CC oligonucleotide (ON) with low adenosine (up to 15%), which targets  
 CC nucleic acids involved in bronchoconstriction, allergies, and/or  
 CC inflammation. The ON can have antiinflammatory, antiallergic,  
 CC anticholinergic, cytotatic and analgesic activities. The compositions are  
 CC useful for the treatment of diseases associated with inflammation,  
 CC impaired airways, including lung disease and diseases whose secondary  
 CC effects affect the lungs of a subject. They can be used for treating  
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,  
 CC impaired respiration, respiratory distress syndrome, pain, cystic  
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive

CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,  
 CC carcinomas, and cancers which may metastasize to the lungs, including  
 CC breast and prostate cancer. The reduction of the adenosine content of the  
 CC ONS reduces side effects. The A-containing ONS break down with the  
 CC release of deoxyadenosine which activates adenosine receptors causing  
 CC bronchoconstriction and inflammation. AAA3313 to AAA3512 represent the  
 CC nucleotide sequences given in the sequence listing from the present  
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185  
 CC sequences are also called SEQ ID NO:1 to 185, but the sequences differ  
 CC from the previously named sequences. SEQ ID NO:11 to 1680 (AAA3233 to  
 CC AAA3392) are specifically claimed ONS from the present invention. N.B.  
 CC Sequences given in the disclosure of the present invention do not match  
 CC up with their corresponding SEQ ID NO: sequences given in the sequence  
 CC listing

XX  
 XX Sequence 1742 BP; 405 A; 492 C; 466 G; 379 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 3; Length 1742;

Best Local Similarity 99.8%; Pred. No. 1.6e-129;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAACCGGGAGTCCCTTTTAAAGCACTTGTCTTGTGTCGCAATCGGCTCTCCCA 60  
 Db 76 ATGAACCGGGAGTCCCTTTTAAAGCACTTGTCTTGTGTCGCAATCGGCTCTCCCA 135  
 Qy 61 GCAGCCACTGAGGAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAATGACACC 120  
 Db 136 GCAGCCACTGAGGAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAATGACACC 195  
 Qy 121 TGTACAGCTTCCAGAAAGAGCATCAATTCGACTGGAAAACTCCACAGATAAG 180  
 Db 196 TGTACAGCTTCCAGAAAGAGCATCAATTCGACTGGAAAACTCCACAGATAAG 255  
 Qy 181 ATTCTGGGAATCGGGCTCTTCTTAACATAAGTCCATCCAGCTGAATGATCGGGCT 240  
 Db 256 ATTCTGGGAATCGGGCTCTTCTTAACATAAGTCCATCCAGCTGAATGATCGGGCT 315  
 Qy 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 300  
 Db 316 GACTCAAGAAAGACCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 375  
 Qy 301 ATAGAAAGCTCAGATCTTACATCTGTGAAGTGAAGCAAGAAAGAGGTGCAATTG 360  
 Db 376 ATAGAAAGCTCAGATCTTACATCTGTGAAGTGAAGCAAGAAAGAGGTGCAATTG 435  
 Qy 361 CTAGTGTTCGATTTGATGTCCTCAACTGACCCCACTGCTTCAGGGGCAAGCTTACC 420  
 Db 436 CTAGTGTTCGATTTGATGTCCTCAACTGACCCCACTGCTTCAGGGGCAAGCTTACC 495  
 Qy 421 CTGACCTTGGAGAGCCCTGATGATAGACCCCTCAGTGCATGTAGAGTCCAAAGGGGT 480  
 Db 496 CTGACCTTGGAGAGCCCTGATGATAGACCCCTCAGTGCATGTAGAGTCCAAAGGGGT 555  
 Qy 481 AAAAATACAGAGGGGGGAGAACCTCTCTCGTGTCTCAGCTGAGCTCCAGATATGTC 540  
 Db 556 AAAAATACAGAGGGGGGAGAACCTCTCTCGTGTCTCAGCTGAGCTCCAGATATGTC 615  
 Qy 541 ACCTGGACATGACATGCTTCTTGCAACCAAGAAAGTGGAGTTCAAAATGACATGTC 600  
 Db 616 ACCTGGACATGACATGCTTCTTGCAACCAAGAAAGTGGAGTTCAAAATGACATGTC 675  
 Qy 601 GTGCTAGCTTCCAGAGAGCTCCGG 626  
 Db 676 GTGCTAGCTTCCAGAGAGCTCCAG 701

RESULT 12

AAF21327 ID AAF21327 standard; DNA; 1742 BP.

XX AAF21327;

DT 14-MAR-2001 (first entry)

XX  
 XX Human low adenosine antisense oligonucleotide related sequence #2894.

XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;  
 XX human; airway disorder; bronchoconstriction; lung inflammation;  
 XX surfactant depletion; respiratory; bronchodilator; antiinflammatory;  
 XX immunosuppressive; antiaesthetic; analgesic; hypotensive; cytostatic;  
 XX respiratory obstruction; pulmonary obstruction; impeded respiration;  
 XX surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;  
 XX respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;  
 XX pulmonary hypertension; emphysema; pulmonary transplantation rejection;  
 XX chronic obstructive pulmonary disease; pulmonary infection; bronchitis;  
 XX cancer; ss.

XX Homo sapiens.

XX W0200062736-A2.

XX 26-OCT-2000.

XX 24-MAR-2000; 2000MO-US008020.

XX 06-APR-1999; 99US-0127958P.

XX (UYEC-) UNIV EAST CAROLINA.

XX (NYCE/) NYCE J W.

XX Nyce JW;

XX WPI; 2000-679539/66.

XX Low adenosine (A) content antisense oligonucleotides which do not trigger  
 XX adenosine receptors during metabolism, useful e.g. for treating cancers  
 XX and respiratory obstructions.

XX Disclosure, Page 1329; 1592pp; English.

XX The present invention describes low adenosine (A) content antisense  
 XX oligonucleotides and compositions (I) comprising them. In the antisense  
 XX oligonucleotides the A is replaced by a 'Universal' or alternative base.  
 XX (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,  
 XX immunosuppressive, antiaesthetic, hypotensive and cytostatic activities.  
 XX The antisense oligonucleotides and (I) can be used to down-regulate the  
 XX expression and/or activity of target polypeptides associated with  
 XX lung/respiratory disorders and malignancies, such as stimulating and  
 XX activating peptide factors and transmitters, transcription factors,  
 XX immunoglobulins and antibodies, antibody receptors, cytokines and  
 XX chemokines, endogenously produced specific and non-specific enzymes,  
 XX binding proteins, adhesion molecules and their receptors, cytokine and  
 XX chemokine receptors, adenosine receptors, bradykinin receptors, central  
 XX nervous system (CNS) and peripheral nervous and non-nervous system  
 XX receptors, CNS and peripheral nervous and non-nervous system peptide  
 XX transmitters, defensins, growth factors, vasoactive peptides and  
 XX receptors, binding proteins and malignancy associated proteins. The  
 XX antisense oligonucleotides may be used in this way to treat disorders  
 XX including respiratory obstruction (especially pulmonary obstruction  
 XX and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or  
 XX surfactant hypoproduction which are associated with a disease or  
 XX condition selected from pulmonary vasoconstriction, inflammation,  
 XX allergies, asthma, impeded respiration, respiratory distress syndrome  
 XX (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary  
 XX hypertension, emphysema, chronic obstructive pulmonary disease (COPD),  
 XX pulmonary transplantation rejection, pulmonary infections, bronchitis,  
 XX and/or cancer. AAF18434 to AAF21543 represent human polynucleotide  
 XX fragments and antisense oligonucleotides used in the exemplification of  
 XX the present invention

XX Sequence 1742 BP; 405 A; 492 C; 466 G; 379 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 3; Length 1742;

Best Local Similarity 99.8%; Pred. No. 1.6e-129; Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTGTGCTGCAACTGGCGCTCCTCCCA 60
DB 76 ATGAACCGGGAGTCCCTTTAGGCACTTGTGCTGCAACTGGCGCTCCTCCCA 135
QY 61 GCAGCCACTCAGGGAAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACGTAGC 120
DB 136 GCAGCCACTCAGGGAAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACGTAGC 195
QY 121 TGTACAGCTTCCCAAGAAAGCATACATTCCTGAGAAAATCTCCACCAAGATTAAG 180
DB 196 TGTACAGCTTCCCAAGAAAGCATACATTCCTGAGAAAATCTCCACCAAGATTAAG 255
QY 181 ATTCTGGAAATCAGGGCTCTCTTTAATTAAGGTCATCCAAAGTGAATGATGCGCT 240
DB 256 ATTCTGGAAATCAGGGCTCTCTTTAATTAAGGTCATCCAAAGTGAATGATGCGCT 315
QY 241 GACTCAAGAAAGCCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 316 GACTCAAGAAAGCCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 375
QY 301 ATGAAGACTCAGATACCTTACATCTGTGTAAGTGAAGACCAAGAGAGAGTGCATTTG 360
DB 376 ATGAAGACTCAGATACCTTACATCTGTGTAAGTGAAGACCAAGAGAGAGTGCATTTG 435
QY 361 CTAGTGTGAGATTGACTGCACTGCACTGCACTGCTTCAAGGGGAGAGCCTGACC 420
DB 436 CTAGTGTGAGATTGACTGCACTGCACTGCACTGCTTCAAGGGGAGAGCCTGACC 495
QY 421 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCCCTGATGCAATGTAGAGTCCAGGGGT 480
DB 496 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCCCTGATGCAATGTAGAGTCCAGGGGT 555
QY 461 AAAAATCATACAGGGGGGAAAGACCTCTCTGCTGTCTCAGCTGAGTCCAGATATGTGAC 540
DB 556 AAAAATCATACAGGGGGGAAAGACCTCTCTGCTGTCTCAGCTGAGTCCAGATATGTGAC 615
QY 541 ACCTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAGTCAAAATAGACATCGTG 600
DB 616 ACCTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAGTCAAAATAGACATCGTG 675
QY 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626
DB 676 GTGCTAGCTTTCAGAAAGGCTCCAG 701

```

## RESULT 13

AA165462 standard; DNA; 1742 BP.

AA165462;

10-DEC-2001 (first entry)

Nucleotide sequence of a human polynucleotide.

Human; receptor; DC-SIGN; dendritic cell; T lymphocyte; HIV; gp120;

C-type lectin; ICAM3; HIV entry; T cell; macrophage; HIV infection; ss.

Homo sapiens.

MO200164752-A2.

07-SEP-2001.

28-FEB-2001; 2001WO-US006322.

02-MAR-2000; 2000US-00517605.

(UNIV ) UNIV NEW YORK STATE.

(UNIV-) UNIV NIJMEGEN.

Litman DR, Kwon D, Van Kooyk Y, Geijtenbeek T;

DR WPI; 2001-602565/68.

an antibody for the treatment or prevention of HIV-infection comprises a

gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding of

DC-SIGN due to concomitant conformational change.

Disclosure; Page 121-122; 131pp; English.

The specification describes an antibody which is specific for an antigenic fragment of gp120. This antigenic fragment binds to DC-SIGN or is exposed upon gp120 binding of DC-SIGN due to concomitant conformational change. DC-SIGN is a receptor that is specifically expressed on dendritic cells and facilitates infection of T lymphocytes with HIV. DC-SIGN is identical to a HIV-1 gp120-binding C-type lectin. DC-SIGN binds ICAM-3 (which is expressed constitutively on T lymphocytes) with high affinity. The antibody of the invention inhibits the trans enhancement of HIV entry into a T cell or macrophage facilitated by dendritic cells. The antibody is useful to treat or prevent HIV infection. The present sequence represents a human polynucleotide, which is used in the course of the invention.

Sequence 1742 BP; 405 A; 491 C; 466 G; 380 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 4; Length 1742;

Best Local Similarity 99.8%; Pred. No. 1.6e-129; Mismatches 1; Indels 0; Gaps 0;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAACCGGGAGTCCCTTTAGGCACTTGTGCTGCAACTGGCGCTCCTCCCA 60

DB 76 ATGAACCGGGAGTCCCTTTAGGCACTTGTGCTGCAACTGGCGCTCCTCCCA 135

QY 61 GCAGCCACTCAGGGAAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACGTAGC 120

DB 136 GCAGCCACTCAGGGAAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACGTAGC 195

QY 121 TGTACAGCTTCCCAAGAAAGCATACATTCCTGAGAAAATCTCCACCAAGATTAAG 180

DB 196 TGTACAGCTTCCCAAGAAAGCATACATTCCTGAGAAAATCTCCACCAAGATTAAG 255

QY 181 ATTCTGGAAATCAGGGCTCTCTTTAATTAAGGTCATCCAAAGTGAATGATGCGCT 240

DB 256 ATTCTGGAAATCAGGGCTCTCTTTAATTAAGGTCATCCAAAGTGAATGATGCGCT 315

QY 241 GACTCAAGAAAGCCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 300

DB 316 GACTCAAGAAAGCCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 375

QY 301 ATGAAGACTCAGATACCTTACATCTGTGTAAGTGAAGACCAAGAGAGAGTGCATTTG 360

DB 376 ATGAAGACTCAGATACCTTACATCTGTGTAAGTGAAGACCAAGAGAGAGTGCATTTG 435

QY 361 CTAGTGTGAGATTGACTGCACTGCACTGCACTGCTTCAAGGGGAGAGCCTGACC 420

DB 436 CTAGTGTGAGATTGACTGCACTGCACTGCACTGCTTCAAGGGGAGAGCCTGACC 495

QY 421 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCCCTGATGCAATGTAGAGTCCAGGGGT 480

DB 496 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCCCTGATGCAATGTAGAGTCCAGGGGT 555

QY 461 AAAAATCATACAGGGGGGAAAGACCTCTCTGCTGTCTCAGCTGAGTCCAGAGTATGTGAC 540

DB 556 AAAAATCATACAGGGGGGAAAGACCTCTCTGCTGTCTCAGCTGAGTATGTGAC 615

QY 541 ACCTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAGTCAAAATAGACATCGTG 600

DB 616 ACCTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAGTCAAAATAGACATCGTG 675

QY 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626

DB 676 GTGCTAGCTTTCAGAAAGGCTCCAG 701

## RESULT 14

AB297021  
 ID AB297021 standard; DNA; 1742 BP.  
 XX  
 AC AB297021;  
 XX  
 DT 17-OCT-2003 (first entry)  
 XX  
 DE Human nucleic acid sequence.  
 XX  
 KW Human; antisense; lung dysfunction; nasal airway dysfunction;  
 antiinflammatory steroid; ubiquinone; antiinflammatory; antiallergic;  
 antiasthmatic; hypotensive; immunosuppressive; cyclostatic; gene therapy;  
 antisense gene therapy; respiratory; lung; adenosine sensitivity;  
 adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;  
 lung inflammation; respiratory disease; de.  
 KM  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200285308-A2.  
 XX  
 PD 31-OCT-2002.  
 XX  
 PF 23-APR-2002; 2002WO-US013135.  
 XX  
 PR 24-APR-2001; 2001US-0286137P.  
 XX  
 PA (EPIC-) EPIGENESIS PHARM INC.  
 XX  
 PI Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;  
 PI Miller S, Tang L, Shahbuddin S;  
 XX  
 DR WPI; 2003-229219/22.  
 XX  
 PT Pharmaceutical composition for treating ailments associated with impaired  
 PT respiration, has oligo(s) antisense to specific gene(s) or its  
 PT corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or  
 PT ubiquinone.  
 XX  
 PS Disclosure; SEQ ID NO 12263; 872bp; English.  
 XX  
 CC The invention relates to a novel pharmaceutical composition, which has a  
 CC first active agent comprising an oligonucleotide antisense to the  
 CC initiation codon, coding region, 5' or 3' end genomic flanking regions,  
 CC 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of  
 CC junctions of genes encoding a polypeptide associated with lung and/or  
 CC nasal airway dysfunction and a second active agent comprising an  
 CC antiinflammatory steroid and ubiquinone. A composition of the invention  
 CC has antiinflammatory, antiallergic, antiasthmatic, hypotensive,  
 CC immunosuppressive, and cyclostatic activity. The composition may have a  
 CC use in antisense gene therapy. The composition is useful for treating or  
 CC preventing a respiratory, lung or malignant disease or condition, also  
 CC for enhancing the prophylactic or therapeutic respiratory effect of an  
 CC antiinflammatory steroid in a subject, for reducing or depleting levels  
 CC of, or reducing sensitivity to adenosine, reducing levels of adenosine  
 CC receptor, producing bronchodilation, increasing levels of ubiquinone or  
 CC lung surfactant in a subject's tissue, or treating bronchoconstriction,  
 CC lung inflammation, lung allergies, or a respiratory disease or condition.  
 CC Note: The sequence data for this patent is not represented in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1742 BP; 405 A; 492 C; 466 G; 379 T; 0 U; 0 Other;  
 XX  
 Query Match 43.4%; Score 624.4; DB 10; Length 1742;  
 Best Local Similarity 99.8%; Pred. No. 1.6e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 XX  
 QY 1 ATGAACCGGGAGAGCCCTTTAGGCACTGCTTGGAGTGCAGAACTGGGCGCTCCTCCCA 60  
 DB 76 ATGAACCGGGAGAGCCCTTTAGGCACTGCTTGGAGTGCAGAACTGGGCGCTCCTCCCA 135  
 QY 61 GCAGCCACTCAGGAGAAAGTGTGCTGGGCAAAAAAGGGGATAGAGTGAAGTGAAC 120  
 1  
 1

DB 136 GCAGCCACTCAGGAGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATAGAGTGAAC 195  
 QY 121 TGTACAGCTTCCAGAAAGAGCATATCAATTCACCTGAGAAAACTCCACCATATPAG 180  
 DB 196 TGTACAGCTTCCAGAAAGAGCATATCAATTCACCTGAGAAAACTCCACCATATPAG 255  
 QY 181 ATTCTGGAAATCAGGGCTCTCTTAACTPAAAGTTCATCAAGCTGAATGATCGGCT 240  
 DB 256 ATTCTGGAAATCAGGGCTCTCTTAACTPAAAGTTCATCAAGCTGAATGATCGGCT 315  
 QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAAACTTCCCTGATCATCAAGAACTTTAAG 300  
 DB 316 GACTCAAGAAAGACCTTTGGGACCAAGAAACTTCCCTGATCATCAAGAACTTTAAG 375  
 QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGGACCAAGAGAGAGTGCATTTG 360  
 DB 376 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGGACCAAGAGAGAGTGCATTTG 435  
 QY 361 CTAGTGTTCGATTTGACCTGCAACTGACACCCACTGCTTCAAGGGCAGACCTGACC 420  
 DB 436 CTAGTGTTCGATTTGACCTGCAACTGACACCCACTGCTTCAAGGGCAGACCTGACC 495  
 QY 421 CTGACCTTGAAGACCCCTCGTAGTAGCCCTCAGTGCATGATGTAGAGTCCAAAGGCT 480  
 DB 496 CTGACCTTGAAGACCCCTCGTAGTAGCCCTCAGTGCATGATGTAGAGTCCAAAGGCT 555  
 QY 481 AAAAACAATACAGGGGGGGAAGACCTCTCCGTCTCAGCTGAGAGCTCCAGATAGTGGC 540  
 DB 556 AAAAACAATACAGGGGGGGAAGACCTCTCCGTCTCAGCTGAGAGCTCCAGATAGTGGC 615  
 QY 541 ACCTGACATGACATGCTCTTGCAGAAACGAGAAAGGTGAGTTCAAAATAGACATCGTG 600  
 DB 616 ACCTGACATGACATGCTCTTGCAGAAACGAGAAAGGTGAGTTCAAAATAGACATCGTG 675  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTTCGG 626  
 DB 676 GTGCTAGCTTTCCAGAAAGGCTTCAG 701  
 RESULT 15  
 ID AD131687  
 ID AD131687 standard; cDNA; 1742 BP.  
 XX  
 AC AD131687;  
 XX  
 DT 17-JUN-2004 (first entry)  
 XX  
 DE Human cDNA #1013.  
 XX  
 KW Human; gene; ss; immunological response; immunopathological condition;  
 Crohn's disease; asthma; ulcerative colitis; hyper eosinophilia;  
 irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;  
 acute monocytic leukemia; antiinflammatory; antiasthmatic; antitumor;  
 osteopathic; antiallergic; antineumatic; cyclostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6607879-B1.  
 XX  
 PD 19-AUG-2003.  
 XX  
 PF 09-FEB-1998; 98US-00023655.  
 XX  
 PR 09-FEB-1998; 98US-00023655.  
 XX  
 PA (INCY-) INCYTE CORP.  
 XX  
 PI Cocks BG, Stuart SG, Sellhammer JF;  
 XX  
 DR WPI; 2003-895307/82.  
 XX  
 PT A composition comprising a plurality of cDNAs, useful for detecting  
 altered expression of genes in an immunological response or for



PT diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma  
PT or osteoarthritis.

XX Claim 1; SEQ ID NO 1013; 50pp; English.

CC The invention relates to a composition comprising a plurality of cDNAs  
CC for detecting the altered expression of genes in an immunological  
CC response. The invention also relates to a method of diagnosing or  
CC monitoring the treatment of an immunopathological condition in a sample,  
CC comprising obtaining nucleic acids from a sample, contacting the nucleic  
CC acids of the sample with an array comprising the plurality of cDNAs under  
CC conditions to form one or more hybridisation complexes, detecting the  
CC hybridisation complexes and comparing the levels of the detected  
CC hybridisation complexes with the level of hybridisation complexes  
CC detected in a non-diseased sample, where an altered level of the detected  
CC hybridisation complexes correlates with the presence of an  
CC immunopathological condition. Also disclosed are an expression profile  
CC comprising a microarray and a plurality of detectable complexes and a  
CC method for identifying a plurality of polynucleotide probes. The cDNAs  
CC are useful as hybridisable array elements in a microarray for monitoring  
CC the expression of target polynucleotides. The microarray can be used in  
CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,  
CC ulcerative colitis, hyperesinophilia, irritable bowel syndrome,  
CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in  
CC identifying agents for the treatment of the diseases. The microarray may  
CC also be used in drug discovery and development, toxicological and  
CC carcinogenicity studies, forensics or pharmacogenomics. The composition  
CC may also be used in purification of a subpopulation of mRNAs, cDNAs or  
CC genomic fragments. This sequence data for this patent did not form part of  
CC the printed specification but was obtained in electronic format directly  
CC from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX Sequence 1742 BP; 405 A; 491 C; 466 G; 380 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 11; Length 1742;

Best Local Similarity 99.8%; Pred. No. 1.6e-129;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
OY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGCTGCACTGGCGCTCTCCCA 60
DB 76 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGCTGCACTGGCGCTCTCCCA 135
OY 61 GCAGCCTCTCAGGGAAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAATGACC 120
DB 136 GCAGCCTCTCAGGGAAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAATGACC 195
OY 121 TGTACAGCTTCCAGAGAAGAGCATACAAATTCACCTGGAAAACTCCAAACAGATAAAG 180
DB 196 TGTACAGCTTCCAGAGAAGAGCATACAAATTCACCTGGAAAACTCCAAACAGATAAAG 255
OY 181 ATTCTGGGAATCAGGGCTCTTTTAATTAAGGTCCATCCAGCTGAATGATCGCGT 240
DB 256 ATTCTGGGAATCAGGGCTCTTTTAATTAAGGTCCATCCAGCTGAATGATCGCGT 315
OY 241 GACTCAAGAAGAGCTTTGGGACCAAGAAATTTCCCTGATCATCAAGAAATCTTAAG 300
DB 316 GACTCAAGAAGAGCTTTGGGACCAAGAAATTTCCCTGATCATCAAGAAATCTTAAG 375
OY 301 ATAGAAGACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 360
DB 376 ATAGAAGACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 435
OY 361 CTAGTGTTCGATTAATCACTCTGACCACTGCTTCAGGGGAGAGCTGACC 420
DB 436 CTAGTGTTCGATTAATCACTCTGACCACTGCTTCAGGGGAGAGCTGACC 495
OY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCACTGCAATGTAGAGTCAAGGGGT 480
DB 496 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCACTGCAATGTAGAGTCAAGGGGT 555
OY 481 AAAAATATACAGGGGGGAGAACCTCTCGTGTCTCAGCTGAGCTCAGGATAGTGGC 540
```

```
DB 556 AAAAATATACAGGGGGGAGAACCTCTCGTGTCTCAGCTGAGCTCAGGATAGTGGC 615
OY 541 ACTTGACATGCACTGTCTTTGCAAGAACCAAGAAAGTGGAGTTCAAAATAGACATCGTG 600
DB 616 ACTTGACATGCACTGTCTTTGCAAGAACCAAGAAAGTGGAGTTCAAAATAGACATCGTG 675
OY 601 GTGCTAGCTTCCAGAAAGGCGCTCCGG 626
DB 676 GTGCTAGCTTCCAGAAAGGCGCTCCAG 701
```

Search completed: November 22, 2004, 11:05:33  
Job time : 760 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:29:56 ; Search time 65.8175 Seconds  
(without alignments)  
511.863 Million cell updates/sec

Title: US-09-936-702-3

Perfect score: 2668

Sequence: 1 MNRGVPRHLLVLTALLP.....GTRLEIKLVPRSGHHHHH 508

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

478139

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTCUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1064	39.9	295	6 5223394-9	Patent No. 5223394
2	1064	39.9	318	6 5223394-11	Patent No. 5223394
3	1064	39.9	458	3 US-09-039-555B-15	Sequence 15, Appl
4	1064	39.9	458	3 US-09-517-605-3	Sequence 3, Appl
5	1062	39.8	458	4 US-10-092-138A-25	Sequence 25, Appl
6	1059	39.7	398	2 US-08-284-391B-29	Sequence 29, Appl
7	1059	39.7	398	4 US-08-394-388A-29	Sequence 29, Appl
8	1059	39.7	402	1 US-08-236-311-1	Sequence 1, Appl
9	1059	39.7	402	3 US-08-457-918-1	Sequence 1, Appl
10	1059	39.7	402	4 US-10-157-408-1	Sequence 1, Appl
11	1059	39.7	457	4 US-08-328-500-9	Sequence 4, Appl
12	1059	39.7	458	3 US-08-466-368-4	Sequence 4, Appl
13	1059	39.7	462	2 US-08-417-495-5	Sequence 5, Appl
14	1059	39.7	462	2 US-08-284-391B-5	Sequence 5, Appl
15	1059	39.7	462	3 US-09-218-950-5	Sequence 5, Appl
16	1059	39.7	462	4 US-08-394-388A-5	Sequence 5, Appl
17	1059	39.7	462	5 PCT-US92-01785-5	Sequence 5, Appl
18	1059	39.7	462	5 PCT-US92-01785-5	Sequence 5, Appl
19	1059	39.7	532	2 US-08-417-495-6	Sequence 6, Appl
20	1059	39.7	532	2 US-08-284-391B-6	Sequence 6, Appl
21	1059	39.7	532	2 US-09-218-950-6	Sequence 6, Appl
22	1059	39.7	532	4 US-08-394-388A-6	Sequence 6, Appl
23	1059	39.7	532	5 PCT-US92-01785-6	Sequence 6, Appl
24	1059	39.7	532	5 PCT-US92-01785-6	Sequence 6, Appl
25	1059	39.7	575	2 US-08-417-495-4	Sequence 4, Appl
26	1059	39.7	575	2 US-08-284-391B-4	Sequence 4, Appl
27	1059	39.7	575	2 US-08-284-391B-4	Sequence 4, Appl

28	1059	39.7	575	3 US-09-218-950-4	Sequence 4, Appl
29	1059	39.7	575	4 US-08-394-388A-4	Sequence 4, Appl
30	1059	39.7	575	5 PCT-US92-01785-4	Sequence 4, Appl
31	1059	39.7	575	5 PCT-US92-01785-4	Sequence 4, Appl
32	1059	39.7	630	4 US-08-472-888A-6	Sequence 6, Appl
33	1056	39.6	458	6 5223394-7	Patent No. 5223394
34	1053	39.5	394	4 US-08-466-368-2	Sequence 2, Appl
35	1053	39.5	394	4 US-08-328-500-2	Sequence 2, Appl
36	1051	39.4	530	3 US-08-477-460B-4	Sequence 4, Appl
37	1051	39.4	530	3 US-08-379-516-4	Sequence 4, Appl
38	1051	39.4	530	3 US-09-329-916-4	Sequence 4, Appl
39	1051	39.4	530	3 US-08-485-372A-4	Sequence 4, Appl
40	1051	39.4	530	3 US-09-409-006A-4	Sequence 4, Appl
41	1051	39.4	530	4 US-08-484-681-4	Sequence 4, Appl
42	1051	39.4	530	4 US-09-766-995-4	Sequence 4, Appl
43	1051	39.4	530	5 PCT-US93-07422-4	Sequence 4, Appl
44	1050	39.4	432	3 US-08-477-460B-2	Sequence 2, Appl
45	1050	39.4	432	3 US-08-379-516-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
5223394-9  
; Patent No. 5223394  
; APPLICANT: WALLNER, BARBARA  
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING  
; LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL  
; LINKAGE SIGNAL SEQUENCE  
; NUMBER OF SEQUENCES: 12  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/335,688  
; FILING DATE: 10-APR-1989  
; SEQ ID NO: 9:  
; LENGTH: 295

Query Match 39.9%; Score 1064; DB 6; Length 295;  
Best Local Similarity 100.0%; Pred. No. 8.36-54;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MNRGVPRHLLVLTALLPAAATGKGVVIGKGGDTVELTCAASQKSIQFMKSNQK 60
DB	1	MNRGVPRHLLVLTALLPAAATGKGVVIGKGGDTVELTCAASQKSIQFMKSNQK 60
QY	61	ILVNGSEFLTKGPSKLNDRASRRSLMDQGNFLLIKLKIEDSDTYICEVDEQKEEVL 120
DB	61	ILVNGSEFLTKGPSKLNDRASRRSLMDQGNFLLIKLKIEDSDTYICEVDEQKEEVL 120
QY	121	LVFGLTANSSTHLLQGOSITLTLESPPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB	121	LVFGLTANSSTHLLQGOSITLTLESPPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY	181	TWTCVTLNOKKVEFKIDIVLAFOKAS 208
DB	181	TWTCVTLNOKKVEFKIDIVLAFOKAS 208

RESULT 2  
5223394-11  
; Patent No. 5223394  
; APPLICANT: WALLNER, BARBARA  
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING  
; LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL  
; LINKAGE SIGNAL SEQUENCE  
; NUMBER OF SEQUENCES: 12  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/335,688  
; FILING DATE: 10-APR-1989  
; SEQ ID NO: 11:  
; LENGTH: 318

Query Match 39.9%; Score 1064; DB 6; Length 318;  
Best Local Similarity 100.0%; Pred. No. 9e-54;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALPAAATGKKVVLGKKGDVLELTCTASQKKSIOFHMKNNOIK 60  
DB 1 MNRGVPFRHLLVQLALPAAATGKKVVLGKKGDVLELTCTASQKKSIOFHMKNNOIK 60  
QY 61 ILGNQGSFLLTKGSPSKLNDRADSRSLWDQGNFPLIINLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLLTKGSPSKLNDRADSRSLWDQGNFPLIINLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGSLTLTLTSPGSSPSVQCSPRGKNIQGGKTLVSQLELDG 180  
DB 121 LVFGLTANSDTHLLQGSLTLTLTSPGSSPSVQCSPRGKNIQGGKTLVSQLELDG 180  
QY 181 TWCTVLQONQKVEFKIDIVVLAFOKAS 208  
DB 181 TWCTVLQONQKVEFKIDIVVLAFOKAS 208

RESULT 3  
US-09-039-555B-15  
; Sequence 15, Application US/09039555B  
; Patent No. 6033856  
; GENERAL INFORMATION:  
; APPLICANT: Koerner, Kathrin  
; APPLICANT: Mueller, Rolf  
; APPLICANT: Sadiacek, Hans-Harald  
; TITLE OF INVENTION: PROMOTER OF THE CDC25B GENE, ITS  
; TITLE OF INVENTION: PREPARATION AND USE  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D. C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/039,555B  
; FILING DATE: 16-MAR-1998  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 19710643.9  
; FILING DATE: 14-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bent, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 016779/0131  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 672-5300  
; TELEFAX: (202) 672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 458 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-039-555B-15

Query Match 39.9%; Score 1064; DB 3; Length 458;  
Best Local Similarity 100.0%; Pred. No. 1.3e-53;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALPAAATGKKVVLGKKGDVLELTCTASQKKSIOFHMKNNOIK 60  
DB 1 MNRGVPFRHLLVQLALPAAATGKKVVLGKKGDVLELTCTASQKKSIOFHMKNNOIK 60  
QY 61 ILGNQGSFLLTKGSPSKLNDRADSRSLWDQGNFPLIINLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLLTKGSPSKLNDRADSRSLWDQGNFPLIINLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGSLTLTLTSPGSSPSVQCSPRGKNIQGGKTLVSQLELDG 180  
DB 121 LVFGLTANSDTHLLQGSLTLTLTSPGSSPSVQCSPRGKNIQGGKTLVSQLELDG 180  
QY 181 TWCTVLQONQKVEFKIDIVVLAFOKAS 208  
DB 181 TWCTVLQONQKVEFKIDIVVLAFOKAS 208

RESULT 4  
US-09-517-605-3  
; Sequence 3, Application US/09517605  
; Patent No. 6391567  
; GENERAL INFORMATION:  
; APPLICANT: Littman, Dan R.  
; APPLICANT: Kwon, Douglas S.  
; APPLICANT: van Kooyk, Yvette  
; APPLICANT: Geijtenbeek, Theo  
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO  
; FILE REFERENCE: 1049-1-017  
; CURRENT APPLICATION NUMBER: US/09/517,605  
; CURRENT FILING DATE: 2000-03-02  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO: 3  
; LENGTH: 458  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-517-605-3

Query Match 39.9%; Score 1064; DB 3; Length 458;  
Best Local Similarity 100.0%; Pred. No. 1.3e-53;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALPAAATGKKVVLGKKGDVLELTCTASQKKSIOFHMKNNOIK 60  
DB 1 MNRGVPFRHLLVQLALPAAATGKKVVLGKKGDVLELTCTASQKKSIOFHMKNNOIK 60  
QY 61 ILGNQGSFLLTKGSPSKLNDRADSRSLWDQGNFPLIINLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLLTKGSPSKLNDRADSRSLWDQGNFPLIINLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGSLTLTLTSPGSSPSVQCSPRGKNIQGGKTLVSQLELDG 180  
DB 121 LVFGLTANSDTHLLQGSLTLTLTSPGSSPSVQCSPRGKNIQGGKTLVSQLELDG 180  
QY 181 TWCTVLQONQKVEFKIDIVVLAFOKAS 208  
DB 181 TWCTVLQONQKVEFKIDIVVLAFOKAS 208

RESULT 5  
US-10-092-138A-25  
; Sequence 25, Application US/10092138A  
; Patent No. 6743630  
; GENERAL INFORMATION:  
; APPLICANT: Sato, Takashi  
; TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON  
; TITLE OF INVENTION: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION  
; FILE REFERENCE: 65823/JPW/PT  
; CURRENT APPLICATION NUMBER: US/10/092,138A  
; CURRENT FILING DATE: 2002-03-06  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 25  
LENGTH: 458  
TYPE: PRT  
ORGANISM: human  
US-10-092-138A-25

Query Match 39.8%; Score 1062; DB 4; Length 458;  
Best Local Similarity 99.5%; Pred. No. 1.7e-53;  
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNRGVPRHLLVLQALLPAATQKRVVLGKKGDVVELTCTASQKKSIOFHMKNSNOIK 60  
DB 1 NNRGVPRHLLVLQALLPAATQKRVVLGKKGDVVELTCTASQKKSIOFHMKNSNOIK 60  
QY 61 ILGNQGSFLTKGPSKLNDRADSRRLMDQGNPPLIKIKIEDSPDYICEVEDQKEEYVL 120  
DB 61 ILGNQGSFLTKGPSKLNDRADSRRLMDQGNPPLIKIKIEDSPDYICEVEDQKEEYVL 120  
QY 121 LVFGLTANSDTHLQGOISLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLQGOISLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLQNKQKVEFKIDIVLAFOKAS 208  
DB 181 TWTCTVLQNKQKVEFKIDIVLAFOKAS 208

## RESULT 6

US-08-284-391B-29  
Sequence 29, Application US/08284391B  
Patent No. 5851828  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Eiding LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391B  
FILING DATE: 02-AUG-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Eiding, Karen L  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 398 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-284-391B-29

Query Match 39.7%; Score 1059; DB 2; Length 398;  
Best Local Similarity 99.5%; Pred. No. 2.2e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNRGVPRHLLVLQALLPAATQKRVVLGKKGDVVELTCTASQKKSIOFHMKNSNOIK 60  
DB 1 NNRGVPRHLLVLQALLPAATQKRVVLGKKGDVVELTCTASQKKSIOFHMKNSNOIK 60  
QY 61 ILGNQGSFLTKGPSKLNDRADSRRLMDQGNPPLIKIKIEDSPDYICEVEDQKEEYVL 120  
DB 61 ILGNQGSFLTKGPSKLNDRADSRRLMDQGNPPLIKIKIEDSPDYICEVEDQKEEYVL 120  
QY 121 LVFGLTANSDTHLQGOISLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLQGOISLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLQNKQKVEFKIDIVLAFOKAS 208  
DB 181 TWTCTVLQNKQKVEFKIDIVLAFOKAS 208

## RESULT 7

US-09-218-950-29  
Sequence 29, Application US/09218950  
Patent No. 6284240  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Eiding LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/218,950  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Eiding, Karen L  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:  
LENGTH: 398 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-218-950-29

Query Match 39.7%; Score 1059; DB 3; Length 398;  
Best Local Similarity 99.5%; Pred. No. 2.2e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVQLALPAAATGKVVLEKKGDVLELTCTASQKSIQFHMNSNQIK 60  
DB 1 MNRGVPFRHLVQLALPAAATGKVVLEKKGDVLELTCTASQKSIQFHMNSNQIK 60  
QY 61 ILNGQSFLLTKGPKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILNGQSFLLTKGPKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDG 180  
DB 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDG 180  
QY 181 TWICTVLQNGKVEFKIDIVLAFQKAS 208  
DB 181 TWICTVLQNGKVEFKIDIVLAFQKAS 208

RESULT 8  
US-08-394-388A-29  
Sequence 29, Application US/08394388A  
Patent No. 6753162

## GENERAL INFORMATION:

APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INJECTED  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/394,388A  
FILING DATE: 24-FEB-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045

TELEX:  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 398 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-394-388A-29

Query Match 39.7%; Score 1059; DB 4; Length 398;  
Best Local Similarity 99.5%; Pred. No. 2.2e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVQLALPAAATGKVVLEKKGDVLELTCTASQKSIQFHMNSNQIK 60  
DB 1 MNRGVPFRHLVQLALPAAATGKVVLEKKGDVLELTCTASQKSIQFHMNSNQIK 60  
QY 61 ILNGQSFLLTKGPKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILNGQSFLLTKGPKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDG 180  
DB 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDG 180  
QY 181 TWICTVLQNGKVEFKIDIVLAFQKAS 208  
DB 181 TWICTVLQNGKVEFKIDIVLAFQKAS 208

RESULT 9  
US-08-236-311-1  
Sequence 1, Application US/08236311  
Patent No. 556335

## GENERAL INFORMATION:

APPLICANT: Capon, Daniel J.  
APPLICANT: Gregory, Timothy J.  
TITLE OF INVENTION: Adhesion Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/236,311  
FILING DATE: 02-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Haack, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 444PIC2  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-236-311-1

Query Match 39.7%; Score 1059; DB 1; Length 402;  
Best Local Similarity 99.5%; Pred. No. 2.2e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALPAATQGGKGVLTCTASQKKSIOFHMKNNOIK 60  
DB 1 MNRGVPRHLLVQLALPAATQGGKGVLTCTASQKKSIOFHMKNNOIK 60  
QY 61 IINGGSPFLTKGPKSLNDRADSRSLMDQGNPFLIIKLIKIEDSDTYICEVEDQKEEYOL 120  
DB 61 IINGGSPFLTKGPKSLNDRADSRSLMDQGNPFLIIKLIKIEDSDTYICEVEDQKEEYOL 120  
QY 121 LVFGLTANSDBTLLOGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELODGG 180  
DB 121 LVFGLTANSDBTLLOGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELODGG 180  
QY 181 TWTCTVLONQKKEFKIDIVLAFOKAS 208  
DB 181 TWTCTVLONQKKEFKIDIVLAFOKAS 208

RESULT 10  
US-08-457-918-1  
Sequence 1, Application US/08457918

GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
TITLE OF INVENTION: Timothy J.  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,918  
FILING DATE: 1-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Kubinec, Jeffrey S.

REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0444P1C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-457-918-1

Query Match 39.7%; Score 1059; DB 3; Length 402;  
Best Local Similarity 99.5%; Pred. No. 2.2e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALPAATQGGKGVLTCTASQKKSIOFHMKNNOIK 60  
DB 1 MNRGVPRHLLVQLALPAATQGGKGVLTCTASQKKSIOFHMKNNOIK 60  
QY 61 IINGGSPFLTKGPKSLNDRADSRSLMDQGNPFLIIKLIKIEDSDTYICEVEDQKEEYOL 120  
DB 61 IINGGSPFLTKGPKSLNDRADSRSLMDQGNPFLIIKLIKIEDSDTYICEVEDQKEEYOL 120  
QY 121 LVFGLTANSDBTLLOGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELODGG 180  
DB 121 LVFGLTANSDBTLLOGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELODGG 180  
QY 181 TWTCTVLONQKKEFKIDIVLAFOKAS 208  
DB 181 TWTCTVLONQKKEFKIDIVLAFOKAS 208

RESULT 11  
US-10-157-408-1  
Sequence 1, Application US/10157408

GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
TITLE OF INVENTION: Gregory, Timothy J.  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/157,408  
FILING DATE: 28-MAY-2002  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,918  
FILING DATE: 1-JUN-1995  
APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:

NAME: Kubinec, Jeffrey S.  
REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0444P1C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-10-157-408-1  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
Query Match 39.7%; Score 1059; DB 4; Length 402;  
Best Local Similarity 99.5%; Pred. No. 2.2e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MNRGVPFRHLVLVQLALPAPATGKKVVLGKGGDTVELTCTASQKKSIOFHMKNNSQIK 60  
DB 1 MNRGVPFRHLVLVQLALPAPATGKNKVLGKGGDTVELTCTASQKKSIOFHMKNNSQIK 60  
QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLWDQGNFPLIIKLIKIEDSDTYICEVEDQKEEVQL 120  
DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLWDQGNFPLIIKLIKIEDSDTYICEVEDQKEEVQL 120  
QY 121 LVFGLTNSDTHLLOQGSLLTLESPPGSSPSVQCRSPRKNIOGKTLVSQLELDG 180  
DB 121 LVFGLTNSDTHLLOQGSLLTLESPPGSSPSVQCRSPRKNIOGKTLVSQLELDG 180  
QY 181 TWICTVLQNKVKEFKIDIVLAFQKAS 208  
DB 181 TWICTVLQNKVKEFKIDIVLAFQKAS 208  
RESULT 12  
US-08-328-500-9  
Sequence 9, Application US/08328500  
Patent No. 6673896  
GENERAL INFORMATION:  
APPLICANT: Maddon, Paul J.  
APPLICANT: Axel, Richard  
APPLICANT: Sweet, Richard W.  
APPLICANT: Arthos, James  
TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/328,500  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0575/24577-CY  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 457 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-328-500-9  
Query Match 39.7%; Score 1059; DB 4; Length 457;  
Best Local Similarity 99.5%; Pred. No. 2.5e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MNRGVPFRHLVLVQLALPAPATGKKVVLGKGGDTVELTCTASQKKSIOFHMKNNSQIK 60  
DB 1 MNRGVPFRHLVLVQLALPAPATGKNKVLGKGGDTVELTCTASQKKSIOFHMKNNSQIK 60  
QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLWDQGNFPLIIKLIKIEDSDTYICEVEDQKEEVQL 120  
DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLWDQGNFPLIIKLIKIEDSDTYICEVEDQKEEVQL 120  
QY 121 LVFGLTNSDTHLLOQGSLLTLESPPGSSPSVQCRSPRKNIOGKTLVSQLELDG 180  
DB 121 LVFGLTNSDTHLLOQGSLLTLESPPGSSPSVQCRSPRKNIOGKTLVSQLELDG 180  
QY 181 TWICTVLQNKVKEFKIDIVLAFQKAS 208  
DB 181 TWICTVLQNKVKEFKIDIVLAFQKAS 208  
RESULT 13  
US-08-466-368-4  
Sequence 4, Application US/08466368  
Patent No. 6093539  
GENERAL INFORMATION:  
APPLICANT: Maddon, Paul J.  
APPLICANT: Littman, Dan R.  
APPLICANT: Chess, Leonard  
APPLICANT: Axel, Richard  
APPLICANT: Weiss, Robin  
APPLICANT: McDougal, J. S.  
TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN  
TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/466,368  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 24577-B1-B/JPM/AKC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-391-0525  
TELEFAX: 212-391-0525  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 458 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-466-368-4  
Query Match 39.7%; Score 1059; DB 3; Length 458;  
Best Local Similarity 99.5%; Pred. No. 2.5e-53;



Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMRGVPFRLHLVLQALLPATQGNKVVLGKKGDVVELTCTASQKKSIOFHMKNQNIK 60  
DB 1 NMRGVPFRLHLVLQALLPATQGNKVVLGKKGDVVELTCTASQKKSIOFHMKNQNIK 60  
QY 61 ILGNQGSFLTGPSPKLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLTGPSPKLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGITANSDDHLLQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGITANSDDHLLQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLONQKKEFKIDIVLAFQKAS 208  
DB 181 TWTCTVLONQKKEFKIDIVLAFQKAS 208

## RESULT 14

US-08-417-495-5  
Sequence 5, Application US/08417495  
Patent No. 5843728  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian et al.  
TITLE OF INVENTION: Redirection of Cellular Immunity by  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 558X  
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
SOFTWARE: Wordperfect (Version 5.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,495  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/203,866  
FILING DATE:  
APPLICATION NUMBER: US/07/847,566  
FILING DATE:  
APPLICATION NUMBER: 07/665,961  
FILING DATE: March 7, 1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/119002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 462 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-417-495-5

Query Match 39.7%; Score 1059; DB 2; Length 462;  
Best Local Similarity 99.5%; Pred. No. 2.5e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 NMRGVPFRLHLVLQALLPATQGNKVVLGKKGDVVELTCTASQKKSIOFHMKNQNIK 60

DB 1 NMRGVPFRLHLVLQALLPATQGNKVVLGKKGDVVELTCTASQKKSIOFHMKNQNIK 60

QY 61 ILGNQGSFLTGPSPKLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLTGPSPKLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120

QY 121 LVFGITANSDDHLLQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGITANSDDHLLQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180

QY 181 TWTCTVLONQKKEFKIDIVLAFQKAS 208  
DB 181 TWTCTVLONQKKEFKIDIVLAFQKAS 208

## RESULT 15

US-08-284-391B-5  
Sequence 5, Application US/08284391B  
Patent No. 5851828  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391B  
FILING DATE: 02-AUG-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 462 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-284-391B-5

Query Match 39.7%; Score 1059; DB 2; Length 462;  
Best Local Similarity 99.5%; Pred. No. 2.5e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 NMRGVPFRLHLVLQALLPATQGNKVVLGKKGDVVELTCTASQKKSIOFHMKNQNIK 60

Db 1 ||||| 60  
QY 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIITKLIKIEDSDTYICEVEDQKEEVOL 120  
Db 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIITKLIKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTFANSDTHLLQGSFLTITLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELOD SG 180  
Db 121 LVFGLTFANSDTHLLQGSFLTITLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELOD SG 180  
QY 181 TWTCTVLQNGKVEFKIDIVVLAFOKAS 208  
Db 181 TWTCTVLQNGKVEFKIDIVVLAFOKAS 208

Search completed: November 21, 2004, 13:44:45  
Job time : 66.8175 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:42:32 : Search time 222.482 Seconds  
(without alignments)  
808.591 Million cell updates/sec

Title: US-09-936-702-3

Perfect score: 2668  
Sequence: 1 MNRGVPRFHLVLVQLALLP.....GTRLEIKLVPRSGHHHHH 508

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1570615 seqs, 354127592 residues 1570615

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10D\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1064	39.9	458	US-10-151-274-3	Sequence 3, Appli
2	1064	39.9	458	US-10-103-597A-39	Sequence 39, Appli
3	1064	39.9	458	US-10-188-444-39	Sequence 39, Appli
4	1064	39.9	458	US-10-207-655-170	Sequence 170, App
5	1062	39.8	458	US-08-681-219-27	Sequence 27, Appli
6	1062	39.8	458	US-09-230-111C-25	Sequence 25, Appli
7	1062	39.8	458	US-10-092-138-25	Sequence 25, Appli
8	1059	39.7	402	US-09-936-537-29	Sequence 29, Appli
9	1059	39.7	402	US-10-157-408-1	Sequence 1, Appli
10	1059	39.7	402	US-10-097-044A-1	Sequence 1, Appli
11	1059	39.7	402	US-10-769-247-1	Sequence 1, Appli
12	1059	39.7	457	US-09-891-119A-9	Sequence 9, Appli
13	1059	39.7	462	US-09-936-537-5	Sequence 5, Appli

14	1059	39.7	462	US-09-243-008-5	Sequence 5, Appli
15	1059	39.7	532	US-09-939-537-6	Sequence 6, Appli
16	1059	39.7	532	US-09-243-008-6	Sequence 6, Appli
17	1059	39.7	575	US-09-939-537-4	Sequence 4, Appli
18	1059	39.7	575	US-09-243-008-4	Sequence 4, Appli
19	1053	39.5	397	US-09-891-119A-2	Sequence 2, Appli
20	1051	39.4	530	US-08-485-163-5	Sequence 5, Appli
21	1051	39.4	530	US-09-766-995-4	Sequence 4, Appli
22	1050	39.4	432	US-08-485-163-3	Sequence 3, Appli
23	1050	39.4	432	US-09-766-995-2	Sequence 2, Appli
24	1048	39.3	310	US-08-485-163-7	Sequence 7, Appli
25	1048	39.3	310	US-09-766-995-6	Sequence 6, Appli
26	1035	38.8	203	US-09-939-537-31	Sequence 31, Appli
27	959.5	36.0	250	US-09-880-748-1952	Sequence 1952, Ap
28	959.5	36.0	250	US-10-293-418-1952	Sequence 1952, Ap
29	957.5	35.9	246	US-09-880-748-1268	Sequence 1268, Ap
30	957.5	35.9	246	US-10-293-418-1268	Sequence 1268, Ap
31	954	35.8	251	US-09-880-748-1238	Sequence 1238, Ap
32	954	35.8	251	US-10-293-418-1238	Sequence 1238, Ap
33	946	35.5	434	US-10-157-408-4	Sequence 4, Appli
34	946	35.5	434	US-10-097-044A-4	Sequence 4, Appli
35	946	35.5	434	US-10-769-247-4	Sequence 4, Appli
36	945.5	35.4	448	US-10-024-329-32	Sequence 32, Appli
37	940	35.2	370	US-09-759-841-6	Sequence 6, Appli
38	929	34.8	621	US-10-768-932A-4	Sequence 4, Appli
39	926	34.7	788	US-10-073-118-26	Sequence 26, Appli
40	923	34.6	590	US-09-934-060A-13	Sequence 13, Appli
41	923	34.6	720	US-09-934-060A-2	Sequence 2, Appli
42	923	34.6	720	US-09-934-060A-4	Sequence 4, Appli
43	917	34.4	450	US-10-768-932A-2	Sequence 2, Appli
44	917	34.4	453	US-10-768-932A-8	Sequence 8, Appli
45	916.5	34.4	476	US-10-768-932A-6	Sequence 6, Appli

#### ALIGNMENTS

RESULT 1  
US-10-151-274-3  
Sequence 3, Application US/10151274  
Publication No. US20030064071A1  
GENERAL INFORMATION:  
APPLICANT: Liltman, Dan R.  
APPLICANT: Kwon, Douglas S.  
APPLICANT: van Kooyk, Yvette  
TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY  
TITLE OF INVENTION: INTO  
TITLE OF INVENTION: CELLS  
FILE REFERENCE: 1049-1-017  
CURRENT APPLICATION NUMBER: US/10/151,274  
CURRENT FILING DATE: 2002-05-20  
PRIOR APPLICATION NUMBER: US/09/517,605  
PRIOR FILING DATE: 2000-03-02  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 458  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-151-274-3

Query Match 39.9%; Score 1064; DB 14; Length 458;  
Best Local Similarity 100.0%; Pred. No. 7.9e-51;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRFHLVLVQLALLPAAATGKKVYVGGKDPVETCTTASOKSIOFHMKNSNOIK 60  
DB 1 MNRGVPRFHLVLVQLALLPAAATGKKVYVGGKDPVETCTTASOKSIOFHMKNSNOIK 60  
QY 61 ILNGSGSLTGPSPKLNDRADSRSLMDQGNFPLIKLKLTEDSDTYICEVEDQKEVQL 120  
DB 61 ILNGSGSLTGPSPKLNDRADSRSLMDQGNFPLIKLKLTEDSDTYICEVEDQKEVQL 120

```

Oy      121  LVFGLTANSDPHLLAQGSLTTLTSPSSPVOCRSPRGKNIOGKTLVSQLELODSG 180
        |||||
Db      121  LVFGLTANSDPHLLAQGSLTTLTSPSSPVOCRSPRGKNIOGKTLVSQLELODSG 180

Oy      181  TWCTVLQNKQKVEFKIDIVLAFQKAS 208
        |||||
Db      181  TWCTVLQNKQKVEFKIDIVLAFQKAS 208

RESULT 2
US-10-103-597A-39
; Sequence 39, Application US/10103597A
; Publication No. US20030096432a1
; GENERAL INFORMATION:
; APPLICANT: Jakobsen, Bent Karsten
; TITLE OF INVENTION: Screening Methods
; FILE REFERENCE: 102286.142
; CURRENT APPLICATION NUMBER: US/10/103,597A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: PCT/GB00/03579
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 9922352.1
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 458
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-103-597A-39

```

```

RESULT 3
US-10-188-444-39
; Sequence 39, Application US/10188444
; Publication No. US20030104635A1
; GENERAL INFORMATION:
; APPLICANT: Jakobsen, Bent Karsten
; TITLE OF INVENTION: Screening Methods
; FILE REFERENCE: 102286.142 (CIP)
; CURRENT APPLICATION NUMBER: US/10/188,444
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: PCT/GB00/03579
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 9923352.1
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens

```

[illegible]

```

RESULT 4
US-10-207-655-170
; Sequence 170, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 170
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-170

Query Match          39.9%; Score 1064; DB 14; Length 458;
Best Local Similarity 100.0%; Pred. No. 7.9e-51;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1  MNRGVPFRLHLVLVQLALLPATQCKRVVLGKGGDTVELTCTASQCKSIQFHWKNSNQIK 60
DB      1  MNRGVPFRLHLVLVQLALLPATQCKRVVLGKGGDTVELTCTASQCKSIQFHWKNSNQIK 60

QY      61  ILNGGSLFTGKPSKLNDRADSRRLMDQGNPFLIKVLKTIEDSDTYICEVEDQKEEYVL 120
DB      61  ILNGGSLFTGKPSKLNDRADSRRLMDQGNPFLIKVLKTIEDSDTYICEVEDQKEEYVL 120

QY      121  LVFGITANSDFHLQGGSLTTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB      121  LVFGITANSDFHLQGGSLTTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180

QY      181  TWTCTVLQNGKKVEFKIDIVVLAIFKAS 208
DB      181  TWTCTVLQNGKKVEFKIDIVVLAIFKAS 208

RESULT 5
US-08-681-219-27
; Sequence 27, Application US/08681219
; Publication No. US20020058607A1
; GENERAL INFORMATION:
; APPLICANT: Takaaki Sato and Junn Yanagisawa
; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN
; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF
; TITLE OF INVENTION: (PDZ/ DHR) DOMAIN AND USES THEREOF
; NUMBER OF SEQUENCES: 35

```



```

;
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,537
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284,391
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
;
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-939-537-29

Query Match          39.7%; Score 1059; DB 10; Length 398;
Best Local Similarity 99.5%; Pred. No. 1.3e-50;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLILVQLALPAPATQGGKGVLGKGDVETLTCTASQKKSIOFHMKNNOIK 60
DB 1 MNRGVPFRHLILVQLALPAPATQGGKGVLGKGDVETLTCTASQKKSIOFHMKNNOIK 60
QY 61 ILNGSGFLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120
DB 61 ILNGSGFLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120
QY 121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVOCRSRPRGKNIQGGKTLISVQLDELDSG 180
DB 121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVOCRSRPRGKNIQGGKTLISVQLDELDSG 180
QY 181 TWTCVTLOKQKVEFKIDIVVLAFOKAS 208
DB 181 TWTCVTLOKQKVEFKIDIVVLAFOKAS 208

RESULT 9
US-10-157-408-1
; Sequence 1, Application US/10157408
; Publication No. US20030104535A1
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
; Gregory, Timothy J.
; TITLE OF INVENTION: Adhesion Variants
```

```

;
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/157,408
; FILING DATE: 28-May-2002
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,918
; FILING DATE: 1-JUN-1995
; APPLICATION NUMBER: 08/236311
; FILING DATE: 02-MAY-1994
; APPLICATION NUMBER: 07/936190
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: 07/842777
; FILING DATE: 18-FEB-1992
; APPLICATION NUMBER: 07/250785
; FILING DATE: 28-SEP-1988
; APPLICATION NUMBER: 07/104329
; FILING DATE: 02-OCT-1987
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0444P1C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8228
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-157-408-1

Query Match          39.7%; Score 1059; DB 14; Length 402;
Best Local Similarity 99.5%; Pred. No. 1.3e-50;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLILVQLALPAPATQGGKGVLGKGDVETLTCTASQKKSIOFHMKNNOIK 60
DB 1 MNRGVPFRHLILVQLALPAPATQGGKGVLGKGDVETLTCTASQKKSIOFHMKNNOIK 60
QY 61 ILNGSGFLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120
DB 61 ILNGSGFLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120
QY 121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVOCRSRPRGKNIQGGKTLISVQLDELDSG 180
DB 121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVOCRSRPRGKNIQGGKTLISVQLDELDSG 180
QY 181 TWTCVTLOKQKVEFKIDIVVLAFOKAS 208
DB 181 TWTCVTLOKQKVEFKIDIVVLAFOKAS 208

RESULT 10
US-10-097-044A-1
; Sequence 1, Application US/10097044A
; Publication No. US20030143320A1
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
```

Gregory, Timothy J.  
TITLE OF INVENTION: Adheseon Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/097,044A  
FILING DATE: 28-May-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,918  
FILING DATE: 1-JUN-1995  
APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Kubinec, Jeffrey S.  
REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0444P1C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-097-044A-1

Query Match 39.7%; Score 1059; DB 14; Length 402;  
Best Local Similarity 99.5%; Pred. No. 1.3e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLIQALLPATQKGVVLGKKGDTVELTCTASQKKSIOFHMKNSNOIK 60  
DB 1 MNRGVPFRHLVLIQALLPATQKGVVLGKKGDTVELTCTASQKKSIOFHMKNSNOIK 60  
QY 61 ILNGQSFLLTGKPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYQL 120  
DB 61 ILNGQSFLLTGKPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYQL 120  
QY 121 LVFGLTANSDTHLLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCIVLQNKQKVEFKIDIVVLAFOKAS 208  
DB 181 TWTCIVLQNKQKVEFKIDIVVLAFOKAS 208

RESULT 11  
US-10-769-247-1  
; Sequence 1, Application US/10769247  
; Publication No. US20040197809A1

GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
Gregory, Timothy J.  
TITLE OF INVENTION: Adheseon Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/769,247  
FILING DATE: 30-Jan-2004  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,918  
FILING DATE: 1-JUN-1995  
APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Kubinec, Jeffrey S.  
REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0444P1C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-769-247-1

Query Match 39.7%; Score 1059; DB 17; Length 402;  
Best Local Similarity 99.5%; Pred. No. 1.3e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLIQALLPATQKGVVLGKKGDTVELTCTASQKKSIOFHMKNSNOIK 60  
DB 1 MNRGVPFRHLVLIQALLPATQKGVVLGKKGDTVELTCTASQKKSIOFHMKNSNOIK 60  
QY 61 ILNGQSFLLTGKPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYQL 120  
DB 61 ILNGQSFLLTGKPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYQL 120  
QY 121 LVFGLTANSDTHLLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCIVLQNKQKVEFKIDIVVLAFOKAS 208  
DB 181 TWTCIVLQNKQKVEFKIDIVVLAFOKAS 208

RESULT 12  
US-09-891-119A-9

Sequence 9, Application US/09891119A  
Publication No. US20040013683A1  
GENERAL INFORMATION:  
APPLICANT: Maddon, Paul J.  
TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4  
FILE REFERENCE: 24577-CY-B  
CURRENT APPLICATION NUMBER: US/09/891.119A  
CURRENT FILING DATE: 2001-06-25  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 9  
LENGTH: 457  
TYPE: PRT  
ORGANISM: human  
US-09-891-119A-9

Query Match 39.7%; Score 1059; DB 11; Length 457;  
Best Local Similarity 99.5%; Pred. No. 1.5e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPAPATQKKVYLGKGGDTVELCTASQKKSIQFHMKNNOIK 60  
Db 1 MNRGVPFRHLVLVQLALPAPATQKKVYLGKGGDTVELCTASQKKSIQFHMKNNOIK 60  
QY 61 ILNQGSPFLTKGSPSKLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120  
Db 61 ILNQGSPFLTKGSPSKLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
Db 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLQNKQKVEFKIDIVLAFOKAS 208  
Db 181 TWTCTVLQNKQKVEFKIDIVLAFOKAS 208

RESULT 13  
US-09-939-537-5  
Sequence 5, Application US/09939357  
Publication No. US20030138410A1  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
Banapour, Babak  
Romeo, Charles  
Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbling LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939.537  
FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/284.391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195.395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847.566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665.961

FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbling, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 462 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-939-537-5

Query Match 39.7%; Score 1059; DB 10; Length 462;  
Best Local Similarity 99.5%; Pred. No. 1.5e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPAPATQKKVYLGKGGDTVELCTASQKKSIQFHMKNNOIK 60  
Db 1 MNRGVPFRHLVLVQLALPAPATQKKVYLGKGGDTVELCTASQKKSIQFHMKNNOIK 60  
QY 61 ILNQGSPFLTKGSPSKLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120  
Db 61 ILNQGSPFLTKGSPSKLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
Db 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLQNKQKVEFKIDIVLAFOKAS 208  
Db 181 TWTCTVLQNKQKVEFKIDIVLAFOKAS 208

RESULT 14  
US-09-243-008-5  
Sequence 5, Application US/09243008  
Publication No. US20040005334A1  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian et al.  
TITLE OF INVENTION: Redirection of Cellular Immunity by  
Receptor Chimeras  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
SOFTWARE: WordPerfect (Version 5.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/243.008  
FILING DATE: 02-Feb-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/394.176  
FILING DATE: SEPTEMBER 11, 1995  
APPLICATION NUMBER: 08/203.866  
FILING DATE: February 28, 1994  
APPLICATION NUMBER: 07/847.566  
FILING DATE: March 6, 1992  
APPLICATION NUMBER: 07/665.961  
FILING DATE: March 7, 1991



ATTORNEY/AGENT INFORMATION:  
NAME: Karen F. Lech, Ph.D  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/270001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 462 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-243-008-5

Query Match 39.7%; Score 1059; DB 11; Length 462;  
Best Local Similarity 99.5%; Pred. No. 1.5e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRGVPFRHLIVLQALIPAAATQGNKVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60  
DB 1 MRGVPFRHLIVLQALIPAAATQGNKVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60  
QY 61 ILNGGSPFLTKGPSKLNDRADSRSLMDQGNPFLIIKNLKIEDSDTYICEVEDQKEEYVL 120  
DB 61 ILNGGSPFLTKGPSKLNDRADSRSLMDQGNPFLIIKNLKIEDSDTYICEVEDQKEEYVL 120  
QY 121 LVFGLTANSDFHLLQGOSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDFHLLQGOSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLONQKKVEFKIDIVLAFOKAS 208  
DB 181 TWTCTVLONQKKVEFKIDIVLAFOKAS 208

RESULT 15  
US-09-939-537-6  
Sequence 6, Application US/09939537  
Publication No. US20030138410A1  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
Banapour, Babak  
Romeo, Charles  
Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,537  
FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992

APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 532 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-939-537-6

Query Match 39.7%; Score 1059; DB 10; Length 532;  
Best Local Similarity 99.5%; Pred. No. 1.7e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRGVPFRHLIVLQALIPAAATQGNKVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60  
DB 1 MRGVPFRHLIVLQALIPAAATQGNKVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60  
QY 61 ILNGGSPFLTKGPSKLNDRADSRSLMDQGNPFLIIKNLKIEDSDTYICEVEDQKEEYVL 120  
DB 61 ILNGGSPFLTKGPSKLNDRADSRSLMDQGNPFLIIKNLKIEDSDTYICEVEDQKEEYVL 120  
QY 121 LVFGLTANSDFHLLQGOSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDFHLLQGOSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLONQKKVEFKIDIVLAFOKAS 208  
DB 181 TWTCTVLONQKKVEFKIDIVLAFOKAS 208

Search completed: November 21, 2004, 14:03:56  
Job time : 227.482 secs

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:25:21 ; Search time 2.58212 Seconds  
(without alignments)  
694.642 Million cell updates/sec

Title: US-09-936-702-1

Perfect score: 28

Sequence: 1 GCGGS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*  
1: geneeqp1980s:\*  
2: geneeqp1990s:\*  
3: geneeqp2000s:\*  
4: geneeqp2001s:\*  
5: geneeqp2002s:\*  
6: geneeqp2003as:\*  
7: geneeqp2003bs:\*  
8: geneeqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	28	100.0	5 2 AAR72707	Aar72707 linker fo
2	28	100.0	5 2 AAR34034	Aar34034 Linking s
3	28	100.0	5 2 AAR95062	Aar95062 scfv spac
4	28	100.0	5 2 AAM17094	Aam17094 Gly(4)-Se
5	28	100.0	5 2 AAM19543	Aam19543 Chimeric
6	28	100.0	5 2 AAY02127	Aay02127 Peptide 1
7	28	100.0	5 2 AAY25357	Aay25357 IFNAR2/IF
8	28	100.0	5 2 AAY33597	Aay33597 VH-VL dom
9	28	100.0	5 2 AAY43496	Aay43496 linker fo
10	28	100.0	5 3 AAY83210	Aay83210 Peptide 1
11	28	100.0	5 3 AAB06226	Aab06226 Expressio
12	28	100.0	5 3 AAY54917	Aay54917 linker fr
13	28	100.0	5 3 AAY43750	Aay43750 linker us
14	28	100.0	5 3 AAR14535	Aar14535 Peptide 1
15	28	100.0	5 3 AAB00156	Aab00156 linker us
16	28	100.0	5 4 AAE06269	Aae06269 Glycine 1
17	28	100.0	5 4 AAE11873	Aae11873 Spacer pe
18	28	100.0	5 4 AAB57374	Aab57374 Flexible
19	28	100.0	5 4 AAB56480	Aab56480 Human sin
20	28	100.0	5 4 AAG62343	Aag62343 linker pe
21	28	100.0	5 4 AAB50848	Aab50848 Fluoresce
22	28	100.0	5 4 AAB57784	Aab57784 Peptide 1
23	28	100.0	5 4 AAB55527	Aab55527 Peptide 1
24	28	100.0	5 4 AAB81033	Aab81033 linker pe
25	28	100.0	5 4 AAG62613	Aag62613 linker pe

26	28	100.0	5 4 AAU05173	Aau05173 Pain-rel
27	28	100.0	5 4 AAU05172	Aau05172 Pain-rel
28	28	100.0	5 4 AAG63011	Aag63011 Peptide 1
29	28	100.0	5 4 AAB68571	Aab68571 Linker pe
30	28	100.0	5 4 AAB74580	Aab74580 Context-d
31	28	100.0	5 4 AAE08730	Aae08730 Peptide 1
32	28	100.0	5 5 AABG31871	Aabg31871 Spacer pe
33	28	100.0	5 5 AAU76043	Aau76043 Synthetic
34	28	100.0	5 5 ABB79971	Abb79971 Linker pe
35	28	100.0	5 5 ABB79970	Abb79970 Linker pe
36	28	100.0	5 5 ABB52872	Abb52872 Flexible
37	28	100.0	5 5 ABB31035	Abb31035 Binding d
38	28	100.0	5 5 ABB05006	Abb05006 VH and VL
39	28	100.0	5 5 ABB48193	Abb48193 Peptide 1
40	28	100.0	5 5 ABB60659	Abb60659 Polyclinum
41	28	100.0	5 5 AAO14387	Aao14387 Linker pe
42	28	100.0	5 5 AAE19822	Aae19822 Linker pe
43	28	100.0	5 5 AAE25959	Aae25959 Linker pe
44	28	100.0	5 5 ABB03945	Abb03945 Peptide 1
45	28	100.0	5 5 AAU75495	Aau75495 Human Gly

## ALIGNMENTS

RESULT 1  
ID AAR72707 standard; peptide; 5 AA.

AC AAR72707;

DT 31-OCT-1995 (first entry)

DE Linker for apo A-I and apo B-100 fusion polypeptide.

KW Apo A-I, LDL cholesterol; low density lipoprotein; fusion polypeptide; linker.

OS Synthetic.

PN US408038-A.

PD 18-APR-1995.

PF 08-OCT-1992; 92US-00959946.

PR 09-OCT-1991; 91US-00774633.

PR 18-JUN-1992; 92US-00901706.

PA (SCRI) SCRIPPS RES INST.

PI Witztum JL, Koduri KR, Young SG, Smith RS, Curtiss LK;

PT WPI; 1993-134378/16.

PT Polypeptide mimic of native apo B-100 and native apo A-I - useful in assays for LDL and HDL in plasma samples.

PS Disclosure; Col 18; 41pp; English.

CC A dispersible apo A-I/B-100 fusion polypeptide is claimed which contains a first AA sequence of apo A-I (see AAR72605) and that includes at least AA sequence positions 120-135 (see AAR72606). The two sequences are operatively linked. An exemplary linking sequence is AAR72707 whose encoding DNA can be ligated between an apo A-I and a B-100 encoding DNA sequence

CC Sequence 5 AA;

Query Match 100.0%; Score 28; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
 |||||  
 Db 1 GGGGS 5

# RESULT 2

AAAR34034  
 ID AAR34034 standard; protein; 5 AA.  
 XX  
 AC AAR34034;

XX  
 DT 25-MAR-2003 (revised)  
 DT 13-AUG-1993 (first entry)

XX  
 DE Linking sequence whose encoding DNA can be ligated between an apo A-I-  
 and a B-100-encoding DNA sequence.

XX  
 KW Lipoprotein; apoprotein; B-100; A-I; LDL; HDL; assay.

XX  
 OS Synthetic.

XX  
 PN WO9307165-A1.

XX  
 PD 15-APR-1993.

XX  
 PF 09-OCT-1992; 92WO-US008634.

XX  
 PR 09-OCT-1991; 91US-00774633.

XX  
 PR 18-JUN-1992; 92US-00901706.

XX  
 PR 08-OCT-1992; 92US-00959946.

XX  
 PA (SCRI ) SCRIPPS RES INST.

XX  
 PI Smith RS, Curtles LK, Koduri KR, Witzum JL, Young SG;

XX  
 DR WPI; 1993-134378/16.

XX  
 PT Polypeptide mimic of native apo B-100 and native apo A-I - useful in  
 assays for LDL and HDL in plasma samples.

XX  
 PS Disclosure; Page 14 and page 35; 137pp; English.

XX  
 CC The inventors claim a portion of the polypeptide contg. apo B-100 that  
 immunoreacts with antibodies secreted by the hybridoma MB47 having ATCC

XX  
 CC Accession No. 8746. Polypeptides specifically claimed include residues  
 CC 217-297, 216-310, 216-331, 216-352, 216-377, 1-377, 205-297, 173-297, 140

XX  
 CC -297. DNA sequences encoding the polypeptides are also claimed. Also  
 CC claimed are a fusion polypeptide that contains: (a) a first amino

XX  
 CC acid residue sequence up to 250 residues in length that includes residues  
 CC 120-135 of apo A-I, (b) a second amino acid residue sequence up to 375

XX  
 CC residues in length that includes residues 217-297 of apo B-100 and DNA  
 CC encoding it. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25

XX  
 CC -MAR-2003 to correct PR field.)

XX  
 SQ Sequence 5 AA;

Query Match 100.0%; Score 28; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
 |||||  
 Db 1 GGGGS 5

# RESULT 3

AAAR5062  
 ID AAR5062 standard; peptide; 5 AA.  
 XX

XX  
 AC AAR5062;

XX  
 DT 18-AUG-1996 (first entry)  
 XX

DE scFv spacer peptide.

XX  
 KW Nucleic acid transfer system; gene transfer; gene therapy;

XX  
 KW cell targeting; multidomain protein; vector; cancer; scfv;

XX  
 KW single chain antibody.

XX  
 OS Synthetic.

XX  
 PN WO9613599-A1.

XX  
 PD 09-MAY-1996.

XX  
 PF 31-OCT-1995; 95WO-EP004270.

XX  
 PR 01-NOV-1994; 94EP-00810627.

XX  
 PA (WELS/) WELS W.

XX  
 PI Wels W, Fominaya J;

XX  
 DR WPI; 1996-239505/24.

XX  
 PT Nucleic acid transfer system for gene therapy, e.g. against cancer -  
 PT includes toxin translocation domain to target nucleic acid to specific  
 cell.

XX  
 PS Disclosure; Page 8; 106pp; English.

XX  
 CC A flexible spacer peptide (AAR5062) is used to link the light chain  
 CC variable domain to the heavy chain variable domain of a single chain

XX  
 CC recombinant antibody (scfv). The scfv may be derived from a monoclonal  
 CC antibody, e.g. Mab FRP5, and forms the ligand domain of a multidomain

XX  
 CC protein (see also AAR5053 and AAR5056-58) that is used with an effector  
 CC nucleic acid in a novel nucleic acid transfer system suitable for gene

XX  
 CC therapy. The ligand domain has a target cell recognition function and  
 CC allows cellular internalization of the multidomain protein/nucleic acid

XX  
 CC complex

XX  
 SQ Sequence 5 AA;

Query Match 100.0%; Score 28; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
 |||||  
 Db 1 GGGGS 5

# RESULT 4

AAAM17094  
 ID AAM17094 standard; peptide; 5 AA.  
 XX

XX  
 AC AAM17094;

XX  
 DT 14-SEP-1999 (first entry)

XX  
 DE Gly(4)-Ser linker peptide for chimeric protein construct.

XX  
 KW Haematopoietic protein; human; granulocyte-colony stimulating factor;

XX  
 KW G-CSF; interleukin; c-mpl ligand; linker; gene therapy; aplastic anaemia;

XX  
 KW stem cell expansion; leukopaenia; neutropaenia; vector; bone marrow;

XX  
 KW thrombocytopenia; blood cell activation; growth.

XX  
 OS Synthetic.

XX  
 PN WO9712985-A2.

XX  
 PD 10-APR-1997.

XX  
 PF 04-OCT-1996; 96WO-US015774.

XX  
 PR 05-OCT-1995; 95US-0004834P.

XX (SEAR ) SEARLE & CO G D.  
 PA  
 XX  
 PI Feng Y, Staten NR, Baum CM, Summers NL, Caparon MH, Bauer SC,  
 PI Zurfluh L, McKeern JP, Klein BK, Lee SC, McWhorter CA, Giri JG;  
 XX  
 DR WPI; 1997-226228/20.  
 XX  
 PT Multi-functional haematopoietic receptor agonists - used to stimulate the  
 PT production of haematopoietic cells in patients.  
 XX  
 PS Disclosure; Page 33; 616pp; English.  
 XX  
 CC The invention relates to a novel haematopoietic protein (HP) comprising  
 CC an amino acid (AA) sequence of formula: R1-L1-R2; R2-L1-R1; R1-R2; or R2-  
 CC R1; where R1 and R2 are independently selected from: (i) a modified human  
 CC granulocyte-colony stimulating factor (hG-CSF) AA sequence; (ii) a  
 CC modified human interleukin-3 (hIL-3) AA sequence; (iii) a modified human  
 CC c-mpl ligand; and a colony stimulating factor (CSF); and L1 = a linker  
 CC capable of linking R1 to R2. This sequence represents an example of a  
 CC linker used to construct the proteins of the invention. Vectors  
 CC comprising the nucleic acid molecules are useful for the recombinant  
 CC production of HP. The nucleic acid molecules are useful in gene therapy.  
 CC The HP's are useful for stimulating the production of haematopoietic  
 CC cells in patients, selective ex vivo expansion of stem cells and for  
 CC treatment of haematopoietic disorders. Disorders that can be treated  
 CC include leukaemia, neutropenia, aplastic anaemia and  
 CC thrombocytopenia. In vitro uses include the ability to stimulate bone  
 CC marrow and blood cell activation and growth before infusion into the  
 CC patients  
 CC  
 SQ Sequence 5 AA;

Query Match 100.0%; Score 28; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 DB 1 GGGGS 5

RESULT 5  
 AAM19543  
 ID AAM19543 standard; peptide; 5 AA.  
 XX  
 AC AAM19543;  
 XX  
 DT 19-FEB-1998 (first entry)  
 XX  
 DE Chimeric protein pentapeptide linker for the MBP moiety and PE moiety.  
 XX  
 KW Pseudomonas exotoxin; myelin basic protein; chimeric protein;  
 KW autoimmune disease; multiple sclerosis; human.  
 XX  
 OS Synthetic.  
 XX  
 FN WO9719179-A1.  
 XX  
 PD 29-MAY-1997.  
 XX  
 PF 17-NOV-1996; 96WO-IL000151.  
 XX  
 PR 17-NOV-1995; 95IL-00116044.  
 PR 26-DEC-1995; 95IL-00116559.  
 XX  
 PA (YISS ) YISSUM RES & DEV CO.  
 XX  
 PI Lorberbaum-Galski H, Steinberger I, Beraud E, Marianovsky I;  
 PI Yarkoni S;  
 XX  
 DR WPI; 1997-298116/27.  
 XX

PT New Pseudomonas exotoxin-myelin basic protein chimeric proteins - used  
 PT for the treatment of auto-immune diseases, particularly multiple  
 PT sclerosis.  
 XX  
 PS Claim 6; Page 29; 54pp; English.  
 XX

CC New chimeric proteins have been developed comprising a Pseudomonas  
 CC aeruginosa exotoxin (PE) moiety linked to a myelin basic protein (MBP)  
 CC moiety selected from: (a) MBP; (b) amino acids 69-88 of guinea-pig MBP or  
 CC an antigenic portion; (c) amino acids 84-102 of human MBP or an antigenic  
 CC portion; (d) amino acids 143-168 of human MBP or an antigenic portion;  
 CC and (e) an amino acid sequence in which one or more amino acids have been  
 CC deleted, added, substituted or mutated in the amino acid sequences of  
 CC (a), (b), (c), or (d), the modified sequences retaining at least 75%  
 CC homology with the amino acid sequences. The present sequence represents  
 CC the preferred pentapeptide linker used to link the MBP moiety and PE  
 CC moiety in a chimeric protein. The chimeric proteins can be used for the  
 CC treatment of autoimmune diseases such as multiple sclerosis. The chimeric  
 CC proteins can specifically target and kill MBP specific T cells while  
 CC having no effect on non-target cells  
 CC  
 SQ Sequence 5 AA;

Query Match 100.0%; Score 28; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 DB 1 GGGGS 5

RESULT 6  
 AAY02127  
 ID AAY02127 standard; protein; 5 AA.  
 XX  
 AC AAY02127;  
 XX  
 DT 16-JUL-1999 (first entry)  
 XX  
 DE Peptide linker used to make multifunctional proteins.  
 XX  
 KW Angiostatin; endostatin; interferon; thrombospondin;  
 KW interferon-inducible protein; platelet factor 4; anti-angiogenic;  
 KW anti-tumor; multifunctional protein; angiogenic-mediated disease; cancer;  
 KW diabetic retinopathy; macular degeneration; arthritis;  
 KW tumor cell production; peptide linker.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO9916889-A1.  
 XX  
 PD 08-APR-1999.  
 XX  
 PF 30-SEP-1998; 98WO-US020464.  
 XX  
 PR 01-OCT-1997; 97US-0060609P.  
 XX  
 PA (SEAR ) SEARLE & CO G D.  
 XX  
 PI Bolanowski MA, Caparon MH, Caspersen GF, Gregory SA, Klein BK;  
 PI McKeern JP;  
 XX  
 DR WPI; 1999-255098/21.  
 XX  
 PT New multifunctional proteins useful for treating angiogenic-mediated  
 PT diseases.  
 XX  
 PS Disclosure; Page 11; 121pp; English.  
 XX  
 CC The specification describes multifunctional proteins which comprise  
 CC combinations of angiostatin, endostatin, interferon, thrombospondin,  
 CC interferon-inducible protein and platelet factor 4, and have anti-

CC angiogenic and/or anti-tumor activity. The multifunctional protein may  
 CC exhibit useful properties such as having similar or greater biological  
 CC activity when compared to a single factor or by having improved half-life  
 CC or decreased adverse side effects, or a combination of these properties.  
 CC The proteins can be used for treating an angiogenic-mediated disease,  
 CC e.g. cancer, diabetic retinopathy, macular degeneration, or arthritis.  
 CC They can also be used for inhibiting the production of tumor cells  
 CC (characteristic of lung, breast, ovarian, prostate, pancreatic, gastric,  
 CC colon, renal, bladder cancers; melanoma, hepatoma, sarcoma and lymphoma)  
 CC in a patient and for inhibiting tumor growth. AAY02125-32 represent  
 CC peptide linkers used to make the multifunctional proteins of the  
 CC invention

XX SQ Sequence 5 AA;

XX QY 1 GGGGS 5  
 XX |||||  
 XX 1 GGGGS 5

XX DB 1 GGGGS 5

XX RESULT 7  
 XX AAY25357  
 XX ID AAY25357 standard; peptide; 5 AA.  
 XX XX  
 XX AAY25357;  
 XX XX  
 XX DT 06-SEP-1999 (first entry)  
 XX XX  
 XX DE IFNAR2/IFN-beta complex peptide fragment 1.  
 XX XX  
 XX KM IFNAR2; IFN-beta; type I interferon; IFNAR/IFN complex; IFN; antiviral;  
 XX KM human interferon alpha/beta receptor; anticancer; immunomodulatory;  
 XX KM anti-arthritis; antidiabetic; treatment; hepatitis; viral infection;  
 XX KM hairy cell leukemia; Kaposi's sarcoma; multiple myeloma; cancer; lupus;  
 XX KM diabetes; multiple sclerosis; rheumatoid arthritis; myasthenia gravis;  
 XX KM acquired immune deficiency syndrome.  
 XX XX  
 XX OS Synthetic.  
 XX XX  
 XX PN WO9932141-A1.  
 XX XX  
 XX PD 01-JUL-1999.  
 XX XX  
 XX PF 18-DEC-1998; 98WO-US026926.  
 XX XX  
 XX PR 19-DEC-1997; 97US-0068295P.  
 XX XX  
 XX PA (ISTF ) ARS APPLIED RES SYSTEMS HOLDING NV.  
 XX (MCIN/) MCINNIS P G.  
 XX XX  
 XX PI Tepper M, Cunningham M, Sherris D, El Tayar N, McKenna S;  
 XX DR WPI; 1999-405115/34.  
 XX XX  
 XX PT Prolonging in vivo activity of type I interferon by complexing.  
 XX XX  
 XX PS Example 8; Page 76; 86pp; English.  
 XX XX  
 CC This invention describes a novel method for prolonging the in vivo effect  
 CC of type I interferon (IFN) by administering IFN as a complex (A) with a  
 CC subunit (I) of the human interferon alpha/beta receptor (IFNAR). The  
 CC product of the invention has antiviral, anticancer, immunomodulatory,  
 CC anti-arthritis and antidiabetic activity. (A) have the antiviral,  
 CC anticancer and immunomodulating activities of IFN, e.g. for treating  
 CC hepatitis and other viral infections, hairy cell leukemia, Kaposi's  
 CC sarcoma, multiple myeloma and other cancers, multiple sclerosis,  
 CC rheumatoid arthritis, myasthenia gravis, diabetes, acquired immune  
 CC deficiency syndrome and lupus. When complexed in (A), the storage life of  
 CC IFN is increased (i.e. it is stabilized against oligomerization, without

CC the need for storage at acidic pH) and its biological effect is  
 CC potentiated

XX SQ Sequence 5 AA;

XX QY 1 GGGGS 5  
 XX |||||  
 XX 1 GGGGS 5

XX DB 1 GGGGS 5

XX RESULT 8  
 XX AAY33597  
 XX ID AAY33597 standard; protein; 5 AA.  
 XX XX  
 XX AAY33597;  
 XX XX  
 XX DT 20-DEC-1999 (first entry)  
 XX XX  
 XX DE VH-VL domain linker peptide #9.  
 XX XX  
 XX KM Antigen binding; single chain; variable domain; VH domain; light chain;  
 XX KM heavy immunoglobulin chain; VL domain; anticancer; antiviral; tumor;  
 XX KM antibacterial; antimetastatic; antiinflammatory; treatment; prevention;  
 XX KM diagnosis; vaccine; autoimmune disease; inflammation; blood disorder;  
 XX KM transplant rejection; arthritis; nervous system disorder; infection.  
 XX XX  
 XX OS Synthetic.  
 XX XX  
 XX PN DE19816141-A1.  
 XX XX  
 XX PD 14-OCT-1999.  
 XX XX  
 XX PF 09-APR-1998; 98DE-01016141.  
 XX XX  
 XX PR 09-APR-1998; 98DE-01016141.  
 XX XX  
 XX PA (HMRI ) HOECHST MARION ROUSSEL DEUT GMBH.  
 XX (Kontermann R, Sedlacek H, Mueller R;  
 XX XX  
 XX PI Kontermann R, Sedlacek H, Mueller R;  
 XX DR WPI; 1999-581511/50.  
 XX XX  
 XX PT New polypeptide binding agents containing variable heavy and light  
 XX PT constructs connected via peptide linker, used for treatment, prevention  
 XX PT or diagnosis of e.g. cancer.  
 XX XX  
 XX PS Claim 7; Page 17; 20pp; German.  
 XX XX  
 CC This sequence represents a novel single-chain molecule (I) that binds  
 CC multiple antigens and comprises two variable domains of heavy  
 CC immunoglobulin chains (VH), having specificities A and B and two variable  
 CC domains of light chains (VL), also with specificities A and B. The  
 CC domains are provided as two VH-VL constructs which are attached via a  
 CC peptide (P). Any VH and VL may be replaced by their functional fragments.  
 CC The products of the invention have anticancer, antiviral, antibacterial,  
 CC antimetastatic and antiinflammatory activity. (I) are used to treat;  
 CC prevent or diagnose tumors (e.g. as tumor vaccines), autoimmune diseases  
 CC and inflammation (e.g. transplant rejection and arthritis), blood  
 CC disorders (e.g. of the coagulation and/or circulatory systems, such as  
 CC anemia, leucopenia, thrombocytopenia and hypertension), nervous system  
 CC disorders and/or infections (by viruses or bacteria, or malaria),  
 CC including, when (I) include a fusogenic peptide, use for gene transfer.  
 CC (I) are produced simply and in predominantly homogeneous form, in a wide  
 CC variety of hosts, either in secreted or membrane-bound forms. This  
 CC sequence represents a VH-VL domain linker peptide which is used to  
 CC illustrate the method of the invention

XX SQ Sequence 5 AA;

Query Match 100.0%; Score 28; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCGS 5  
 DB 1 GGCGS 5

RESULT 9  
 AAY43496  
 ID AAY43496 standard; peptide; 5 AA.  
 AC AAY43496;  
 XX  
 DT 26-JAN-2000 (first entry)  
 DE  
 XX

Linker for dual avb3 receptor/metastasis-associated receptor ligands.

KW Interferon-alpha-2b; IFN-alpha; avb3 antagonist; avb3 receptor ligand;  
 KW metastasis-associated receptor ligand; angiogenesis; cell proliferation;  
 KW anti-angiogenic protein; avb3-integrin; cancer; arthritis;  
 KW macular degeneration; diabetic retinopathy; hemangioma; psoriasis;  
 KW osteoporosis; thrombosis; angina; atherosclerosis; antiviral;  
 KW antibacterial; antifungal.

OS Homo sapiens.  
 XX  
 PN WO9951638-A1.  
 XX  
 PD 14-OCT-1999.  
 XX  
 PF 07-APR-1999; 99WO-US004295.  
 XX  
 PR 08-APR-1998; 98US-0081074P.  
 XX  
 PA (SEAR ) SEARLE & CO G D.  
 XX  
 PI Tjoeng FS, Fok KF;  
 XX  
 DR WPI; 1999-620196/53.

PT New conjugates of integrin antagonist and ligand for metastasis-  
 PT associated receptor, for treating angiogenesis-related diseases, e.g.  
 PT cancer.

Claim 18; Page 86; 108pp; English.

CC The present sequence represents a linker used to join the avb3 antagonist  
 CC and the metastasis-associated receptor ligand, in the pharmaceutical  
 CC compounds of the invention. These compounds are dual avb3  
 CC receptor/metastasis-associated receptor ligands, and inhibit angiogenesis  
 CC and thus proliferation of (cancer) cells. One component binds to the avb3  
 CC receptor and the other to a metastasis-associated receptor. The avb3  
 CC antagonists may also be conjugated to anti-angiogenic proteins, such as  
 CC IFN-alpha and its derivatives. The compounds are used to treat  
 CC angiogenesis-related disorders (mediated by the avb3-integrin),  
 CC specifically cancer (of lung, breast, ovary, prostate, stomach, colon,  
 CC kidney or bladder, also melanoma, hepatoma, sarcoma and lymphoma),  
 CC arthritis and macular degeneration, and also diabetic retinopathy,  
 CC hemangioma, psoriasis, osteoporosis, thrombosis, angina, atherosclerosis  
 CC etc. The compounds may also be useful as antiviral, antibacterial and  
 CC antifungal agents

CC  
 XX  
 SQ Sequence 5 AA;

Query Match 100.0%; Score 28; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCGS 5  
 DB 1 GGCGS 5

RESULT 10  
 AAY83210  
 ID AAY83210 standard; peptide; 5 AA.  
 XX  
 AC AAY83210;  
 XX  
 DT 24-JUL-2000 (first entry)  
 DE  
 XX

Peptide linker used in construction of a\_vb\_3 integrin/IFN alpha.

KW Biconjugate: a\_vb\_3 integrin; interferon alpha; angiogenesis; cancer;  
 KW tumour; osteoporosis; Paget's disease; Karpov's sarcoma;  
 KW periodontal disease; metastasis; neoplasia; retinopathy; arthritis;  
 KW psoriasis; leukaemia; malignant melanoma; atherosclerosis;  
 KW smooth muscle cell migration; inhibition; treatment; antagonist; angina;  
 KW thrombosis; restenosis; antiviral; antifungal; antibacterial.

OS Synthetic.  
 XX  
 PN WO200009143-A1.  
 XX  
 PD 24-FEB-2000.  
 XX  
 PF 07-APR-1999; 99WO-US004296.  
 XX  
 PR 13-AUG-1998; 98US-0096442P.  
 XX  
 PA (SEAR ) SEARLE & CO G D.  
 XX  
 PI Fok KF, Tjoeng FS;  
 XX  
 DR WPI; 2000-205894/18.

PT New biconjugates comprising an avb3 antagonist and a metastatic-  
 PT associated receptor ligand, useful for treating cancer and other  
 PT angiogenic diseases, or as antiviral, antifungal or antibacterial agents.

Claim 19; Page 88; 123pp; English.

CC Biconjugates comprising one or more a\_vb\_3 antagonist moieties coupled  
 CC to a peptide or polypeptide having anti-angiogenic properties can be used  
 CC for treating a human patient with an angiogenesis-mediated disease, e.g.  
 CC cancer, arthritis, or macular degeneration. The a\_vb\_3 integrin is  
 CC normally associated with endothelial cells but can promote the formation  
 CC of blood vessels (angiogenesis) in tumours. The a\_vb\_3 integrin is also  
 CC known to play a role in tumour metastasis, neoplasia, osteoporosis,  
 CC Paget's disease, retinopathy, arthritis, periodontal disease, psoriasis  
 CC and smooth muscle cell migration. Interferon alpha is a family of  
 CC proteins which possess complex antiviral, antineoplastic and  
 CC immunomodulating activities. Interferon alpha is effective against a  
 CC variety of cancers including hairy cell leukaemia, chronic myelogenous  
 CC leukaemia, malignant melanoma and Karpov's sarcoma. Multi-functional  
 CC biconjugates comprising both a\_vb\_3 antagonists and interferon alpha 2b  
 CC can exhibit greater biological activity when compared to a single factor  
 CC or having improved half-life or decreased adverse side effects, or a  
 CC combination of these properties. They can be used for inhibiting elevated  
 CC levels of tumor antigens, inhibiting the proliferation of tumor cells and  
 CC inhibiting tumor growth. The biconjugates can also be used for treating  
 CC e.g. osteoporosis, humoral hypercalcemia of malignancy, Paget's disease,  
 CC retinopathy including diabetic retinopathy, arthritis, including  
 CC rheumatoid arthritis, periodontal disease, psoriasis, thrombosis, angina,  
 CC atherosclerosis, smooth muscle cell migration and restenosis in a mammal.  
 CC They are also useful as antiviral, antifungal and antibacterial agents.  
 CC This sequence is a peptide linker used in the construction of the multi-  
 CC functional biconjugates

CC  
 XX  
 SQ Sequence 5 AA;

Query Match 100.0%; Score 28; DB 3; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5  
 |||||  
 DB 1 GGGGS 5

## RESULT 11

AA06226 standard; peptide; 5 AA.

XX AAB06226;

DT 22-NOV-2000 (first entry)

DE Expression vector CANTAB5E inserted peptide.

XX Modified RNase; eosinophil derived neurotoxin protein; EDN; cancer;

KM Kaposi's sarcoma; neoplastic endothelial cell;

XX non-neoplastic endothelial cell; expression vector.

OS Synthetic.

PN W0200026233-A1.

PD 11-MAY-2000.

PF 01-NOV-1999; 99WO-US025737.

PR 02-NOV-1998; 98US-0106732P.

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PI Rybak SM, Newton DL;

DR WPI; 2000-365565/31.

PT N-terminally modified RNase A targeted to and are cytotoxic to cancerous

PT endothelial cells used to treat especially Kaposi's sarcoma.

XX Example 9; Page 34; 51pp; English.

XX The present sequence is a peptide which was inserted into expression

CC vector pCANTAB5E to enable more flexible folding of the human eosinophil

CC derived neurotoxin protein (EDN), which was expressed by the vector. This

CC protein can be directed to cancerous cells using additional N-terminal

CC peptides, where it exerts a cytotoxic effect. The protein can, therefore,

CC be used to treat cancer, particularly Kaposi's sarcoma, and to

CC selectively kill neoplastic and non-neoplastic endothelial cells

XX Sequence 5 AA;

Query Match 100.0%; Score 28; DB 3; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.7e+06;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5  
 |||||  
 DB 1 GGGGS 5

RESULT 12

AA54917 standard; peptide; 5 AA.

XX AAY54917;

DT 14-FEB-2000 (first entry)

DE Linker from IL-12 fusion protein.

XX Interleukin-12; IL-12; fusion protein; IL-12 p35 subunit; B7 protein;

KM IL-12 p40 subunit; gene therapy; tumour; leukaemia; linker.

OS Synthetic.

XX US5994104-A.

XX 30-NOV-1999.

PF 08-NOV-1996; 96US-00751767.

PR 08-NOV-1996; 96US-00751767.

PA (UNLO ) ROYAL FREE HOSPITAL SCHOOL MED.

PI Anderson RJ, Prentice HG, Macdonald ID;

DR WPI; 2000-038261/03.

PT Nucleic acid constructs encoding interleukin-12 fusion proteins useful

PT for treating leukemia and other cancers.

PS Claim 3; Col 93; 73pp; English.

XX This sequence represents a linker that can be used in an interleukin-12

CC fusion protein. The invention relates to an isolated nucleic acid

CC construct (1) comprising a region encoding an interleukin-12 (IL-12)

CC fusion protein (comprising an IL-12 p35 subunit, an IL-12 p40 subunit and

CC a linker peptide (joining the subunits)) and a region encoding a B7

CC protein. (1) may be used to produce IL-12 fusion proteins according to

CC standard recombinant DNA methodologies. The fusion proteins may be

CC produced either in vitro in a fermentation culture or in vivo as part of

CC a gene therapy protocol (in this case (1) is used to transform a patient's

CC cells, which then secrete the functional polypeptide to supplement the

CC patient's own production of IL-12 or to rectify mutations which lead to

CC the expression of inactive polypeptides). The fusion proteins produced in

CC this way may be used to treat any disease which responds to IL-12 such as

CC tumours (both solid and dispersed of the kidney, breast, colon, ovarian

CC and cervical tumours and melanomas) and in particular, tumours of the

CC blood such as leukaemia. Alternatively, the polypeptides may be used as

CC agonists and antagonists of its activity. The antibodies and antagonists

CC may be used to inhibit the activity of IL-12. (1) may also be used

CC diagnostically as a probe which hybridizes to sequences encoding IL-12

CC and the antibodies may be used to detect the presence of IL-12

CC polypeptides in samples. They may be used diagnostically to quantitate

CC the expression of the polypeptide by patients and hence which subjects

CC may be in need of restorative therapy

XX Sequence 5 AA;

Query Match 100.0%; Score 28; DB 3; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.7e+06;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5  
 |||||  
 DB 1 GGGGS 5

RESULT 13

AA43750 standard; peptide; 5 AA.

XX AAY43750;

DT 11-FEB-2000 (first entry)

DE Linker used to construct a bispecific single-chain antibody.

XX bisCD19xCD3 antibody; bispecific single-chain fragment; CD19 antigen;

KM CD3 antigen; CD19-positive target cell; T-cell stimulation;

KM cytotoxic T-lymphocyte; B-cell malignancy; myasthenia gravis;

KM B-cell mediated autoimmune disease; Morbus Basedow;

KM Hashimoto thyroiditis; Goodpasture syndrome; B-cell depletion;

KM non-Hodgkin lymphoma; gene therapy; cancer; viral disease.



[illegible]

PD	13-JUL-2000.
XX	
XX	
PF	10-JAN-2000; 2000WO-US000456.
XX	
PR	08-JAN-1999; 99US-0115404P.
XX	
PR	07-JAN-2000; 2000US-00480336.
XX	
PA	(WILD/) WILD C T.
XX	
PA	(WEIS/) WEISS C D.
XX	
PI	Wild CT, Weiss CD;
XX	
PI	WPI: 2000-465959/40.
XX	
PT	Raising neutralising antibody response to human immunodeficiency virus,
PT	comprises administering a polypeptide capable of forming a stable coiled-
PT	coil solution structure.
XX	
PS	Disclosure; Page 15; 97pp; English.
XX	
XX	
CC	The invention relates to raising a neutralising antibody response to a
CC	broad spectrum of HIV (human immunodeficiency virus) strains and
CC	isolates, comprising the administration of a peptide which corresponds to
CC	or mimics highly conserved portions of the gp41 envelope glycoprotein
CC	which are important in mediating the process of viral entry into host
CC	cells. Such peptides can correspond to or mimic the coiled coil solution
CC	structure of the N-helical domain (the heptad repeat region), or can
CC	correspond or mimic the C-helical domain (the transmembrane-proximal
CC	amphipathic alpha-helical segment), or the gp41 core 6-helix bundle,
CC	which is formed by the interaction of the N- and C-helical domains of
CC	a three gp41 proteins. The peptides can be administered either singly or as
CC	a combination (particularly a combination of N-helical and C-helical
CC	peptides), and can be multimerised. For example, N- and C-helical domain
CC	peptides can be alternately linked together to form a peptide which
CC	mimics the core 6-helix bundle. Administration of the peptide(s)
CC	generates a humoral response, with the production of antibodies against
CC	gp41 structures involved in viral entry. As these portions of gp41 are
CC	well conserved, such antibodies may be effective against a broad range of
CC	HIV strains and isolates. The peptide compositions may be administered as
CC	a prophylactic or therapeutic vaccine to generate antibodies which reduce
CC	or inhibit the ability of HIV to infect uninfected cells. A composition
CC	comprising polyclonal or monoclonal antibodies can be administered to
CC	reduce HIV infection of uninfected cells. Antibodies raised against entry
CC	-relevant gp41 structures may also be used therapeutically and as tools
CC	to further elucidate the mechanism of HIV cell entry. The present
CC	sequence represents a peptide linker which may be used to join peptides
CC	of the invention together to form multimers
XX	
XX	
SO	Sequence 5 AA;
Query Match	100.0%; Score 28; DB 3; Length 5;
Best Local Similarity	100.0%; Pred. NO. 1.7e+06;
Matches 5; Conservative	0; Mismatches 0; Indels 0; Gaps 0.
OY	1 GGGGS 5
DB	1 GGGGS 5
RESULT 15	
AAAB00156	
ID	AAAB00156 standard; peptide; 5 AA.
XX	
AC	AAAB00156;
XX	
DT	08-FEB-2001 (first entry)
XX	
DE	Linker used in sCD4-SCFv(17b) fusion protein.
XX	
KW	Fusion protein; HIV; human immunodeficiency virus; antibody; Fv; AIDS;
KW	acquired immune deficiency syndrome; neutralisation; infection;
KW	gene therapy; CD4; gp120; glycoprotein; resistance; vaccination;
KW	binding domain; single chain antibody; chimera; chimeric protein.

```

XX      Synthetic.
OS
XX      WO200055207-A1.
PN
XX      21-SEP-2000.
PD
XX      16-MAR-2000; 2000WO-US006946.
PF
XX      16-MAR-1999; 99US-0124681P.
PR
XX      (USSH ) US NAT INST OF HEALTH.
PA
XX      Berger EA, Del Castillo CM;
PI
XX      WPI; 2000-638183/61.
DR
XX      Novel neutralizing bispecific fusion proteins effective in viral such as
PT      HIV neutralization, comprises two different binding domains, inducing-
PT      binding domain and induced-binding domain functionally linked by linker.
XX
XX      Claim 30; Page 45; 55pp; English.
PS
XX
XX      bCD4-SCFv(17b) is a neutralizing bispecific fusion protein capable of
CC      binding to two sites of its target protein. The protein comprises a first
CC      binding domain capable of binding to an inducing site on the target
CC      protein, a second binding domain capable of forming neutralising complex
CC      with an induced epitope of the target protein and a linker connecting the
CC      binding domains. bCD4-SCFv(17b) comprises a soluble CD4 fragment
CC      (containing domains D1 and D2) fused to a single chain Fv portion of
CC      antibody 17b via a linker. bCD4-SCFv(17b), its variant, analogue or
CC      mimetic is used for inactivating gp120 protein of HIV, and for
CC      neutralizing HIV. It is also used for blocking and preventing the binding
CC      of the viral or recombinant gp120 protein to soluble CD4 or lymphocyte
CC      CD4 and for inhibiting HIV replication. The chimeric proteins is
CC      therefore useful for treating HIV infection and also AIDS. It is are
CC      particularly useful in the prevention of infection during or immediately
CC      after HIV exposure (e.g., mother/infant transmission, post-exposure
CC      prophylaxis, and as a topical inhibitor) and for providing long term
CC      resistance to HIV infections and AIDS. Gene therapy is used to secrete
CC      the bispecific protein at mucosal surfaces, such as the vaginal, rectal
CC      or oral mucosa. The fusion proteins is highly potent, broadly cross-
CC      reactive with neutralising antibody with high in vivo activity and no FC-
CC      mediated undesirable targeting properties. When the fusion protein is
CC      substantially derived from human proteins, it has minimal immunogenicity
CC      and toxicity in humans which is of great value in prevention of infection
CC      during or immediately after HIV exposure
XX
XX      Sequence 5 AA;
SQ

```

```

Query Match          100.0%; Score 28; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. NO. 1.7e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GGGGS 5
        |||||
Db      1 GGGGS 5

```

Search completed: November 21, 2004, 13:37:11  
 Job time : 5.58212 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:29:56 ; Search time 0.64781 Seconds  
(without alignments)  
511.863 Million cell updates/sec

Title: US-09-936-702-1  
Perfect score: 28  
Sequence: 1 GCGGS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/5C\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/5D\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	28	100.0	5 1 US-07-959-946-6	Sequence 6, Appl
2	28	100.0	5 1 US-08-176-500-140	Sequence 140, App
3	28	100.0	5 1 US-08-471-052A-140	Sequence 140, App
4	28	100.0	5 1 US-08-225-224-54	Sequence 54, Appl
5	28	100.0	5 1 US-08-236-918A-18	Sequence 18, Appl
6	28	100.0	5 1 US-08-463-163-1	Sequence 1, Appl
7	28	100.0	5 1 US-08-566-800A-58	Sequence 58, Appl
8	28	100.0	5 1 US-08-244-469-5	Sequence 5, Appl
9	28	100.0	5 1 US-08-189-331-140	Sequence 140, App
10	28	100.0	5 1 US-08-333-577-6	Sequence 6, Appl
11	28	100.0	5 1 US-08-575-361A-32	Sequence 32, Appl
12	28	100.0	5 2 US-08-564-955-64	Sequence 64, Appl
13	28	100.0	5 2 US-08-528-523-13	Sequence 13, Appl
14	28	100.0	5 2 US-08-537-874-62	Sequence 62, Appl
15	28	100.0	5 2 US-08-448-418-86	Sequence 86, Appl
16	28	100.0	5 2 US-08-471-939-140	Sequence 140, App
17	28	100.0	5 2 US-08-471-800-140	Sequence 140, App
18	28	100.0	5 2 US-08-932-589-58	Sequence 58, Appl
19	28	100.0	5 2 US-08-647-449-26	Sequence 26, Appl
20	28	100.0	5 2 US-08-821-840-1	Sequence 1, Appl
21	28	100.0	5 2 US-08-902-623-75	Sequence 75, Appl
22	28	100.0	5 2 US-08-471-068-140	Sequence 140, App
23	28	100.0	5 2 US-08-308-494A-1	Sequence 1, Appl
24	28	100.0	5 2 US-08-665-202-25	Sequence 25, Appl
25	28	100.0	5 2 US-08-809-668-12	Sequence 12, Appl
26	28	100.0	5 2 US-08-792-553-15	Sequence 15, Appl
27	28	100.0	5 2 US-08-751-767A-24	Sequence 24, Appl

28	28	100.0	5 2 US-08-818-253-40	Sequence 40, Appl
29	28	100.0	5 3 US-08-722-258-54	Sequence 54, Appl
30	28	100.0	5 3 US-09-397-951-12	Sequence 12, Appl
31	28	100.0	5 3 US-08-776-271-3	Sequence 3, Appl
32	28	100.0	5 3 US-08-621-859-64	Sequence 64, Appl
33	28	100.0	5 3 US-09-100-856A-62	Sequence 62, Appl
34	28	100.0	5 3 US-09-046-992-5	Sequence 5, Appl
35	28	100.0	5 3 US-09-215-035-3	Sequence 3, Appl
36	28	100.0	5 3 US-09-075-511-64	Sequence 64, Appl
37	28	100.0	5 3 US-09-099-015-64	Sequence 64, Appl
38	28	100.0	5 3 US-08-818-252-40	Sequence 40, Appl
39	28	100.0	5 3 US-09-091-814-1	Sequence 1, Appl
40	28	100.0	5 3 US-09-281-792B-26	Sequence 26, Appl
41	28	100.0	5 3 US-09-232-863-62	Sequence 62, Appl
42	28	100.0	5 3 US-09-362-805-1	Sequence 1, Appl
43	28	100.0	5 3 US-09-362-805-2	Sequence 2, Appl
44	28	100.0	5 3 US-09-367-953B-26	Sequence 26, Appl
45	28	100.0	5 3 US-09-133-508A-62	Sequence 62, Appl

## ALIGNMENTS

RESULT 1  
US-07-959-946-6  
; Sequence 6, Application US/07959946  
; Patent No. 5408038  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Richard K.  
; APPLICANT: Koduri, Raju  
; APPLICANT: Young, Stephen G.  
; APPLICANT: Wilzium, Joseph L.  
; APPLICANT: Curtiss, Linda K.  
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Goldsmith, Shore, Sucker &  
; ADDRESSEE: Milnamow, Ltd.  
; STREET: 180 No. 5408038th Stetson, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/959,946  
; CLASSIFICATION: 435  
; FILING DATE: 19921008  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,706  
; FILING DATE: 18-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gamsen, Edward P.  
; REGISTRATION NUMBER: 29,381  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312)616-5400  
; TELEFAX: (312)616-5460  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-959-946-6  
Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
Db 1 GGGGS 5

## RESULT 2

US-08-176-500-140  
; Sequence 140, Application US/08176500  
; Patent No. 5498538  
; GENERAL INFORMATION:  
; APPLICANT: Kay, B. K.  
; APPLICANT: Fowles, D. M.  
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
; NUMBER OF SEQUENCES: 141  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/176,500  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/013,416  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mirock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-143  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790-9090  
; TELEFAX: 212 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 140:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-08-176-500-140

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
Db 1 GGGGS 5

## RESULT 3

US-08-471-052A-140  
; Sequence 140, Application US/08471052A  
; Patent No. 5625033  
; GENERAL INFORMATION:  
; APPLICANT: Kay, B. K.  
; APPLICANT: Fowles, D. M.  
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
; NUMBER OF SEQUENCES: 166  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas

CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/471,052A  
; FILING DATE: 06-JUNE-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mirock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-179  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790-9090  
; TELEFAX: 212 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 140:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-08-471-052A-140

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
Db 1 GGGGS 5

## RESULT 4

US-08-225-224-54  
; Sequence 54, Application US/08225224  
; Patent No. 5635599  
; GENERAL INFORMATION:  
; APPLICANT: PASTAN, Ira  
; APPLICANT: KREITMAN, Robert J.  
; TITLE OF INVENTION: CIRCULARLY PERMUTATED LIGANDS AND  
; TITLE OF INVENTION: CIRCULARLY PERMUTED FUSION PROTEINS  
; NUMBER OF SEQUENCES: 57  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Hourie and Crew  
; STREET: Stewart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/225,224  
; FILING DATE: 8-APR-1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen L.  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 15280-193  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 54:

SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-225-224-54

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
|||  
Db 1 GGGGS 5

RESULT 5  
US-08-236-918A-18  
Sequence 18, Application US/08236918A  
Patent No. 5674704  
GENERAL INFORMATION:  
APPLICANT: Alderson, Mark R.  
APPLICANT: Goodwin, Raymond G.  
APPLICANT: Smith, Craig A.  
TITLE OF INVENTION: Cytokine Designated 4-1BB Ligand  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kathryn A. Anderson, Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple 7.5.3  
SOFTWARE: Microsoft Word, Version #6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/236,918A  
FILING DATE: 06-May-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/060,843  
FILING DATE: 07-May-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Anderson, Kathryn A.  
REGISTRATION NUMBER: 32,172  
REFERENCE/DOCKET NUMBER: 2801-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 233-0640  
TELEFAX: (206) 587-0430  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-236-918A-18

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
|||  
Db 1 GGGGS 5

RESULT 6  
US-08-463-163-1

Sequence 1, Application US/08463163  
Patent No. 5696237  
GENERAL INFORMATION:  
APPLICANT: FitzGerald, David J.  
APPLICANT: Chaudhary, Vijay K.  
APPLICANT: Pastan, Ira H.  
APPLICANT: Waldmann, Thomas A.  
APPLICANT: Queen, Cary L.  
TITLE OF INVENTION: Recombinant Antibody-Toxin Fusion Protein  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Street Tower  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105-1492  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,163  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/227,227  
FILING DATE: 22-JAN-1981  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/911,227  
FILING DATE: 24-SEP-1986  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/341,361  
FILING DATE: 21-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/865,722  
FILING DATE: 08-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Ellen L.  
REGISTRATION NUMBER: 32,762  
REFERENCE/DOCKET NUMBER: 015280-12211  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-463-163-1

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
|||  
Db 1 GGGGS 5

RESULT 7  
US-08-566-800A-58  
Sequence 58, Application US/08566800A  
Patent No. 5736364  
GENERAL INFORMATION:  
APPLICANT: Kelley, Robert F.  
APPLICANT: Lazarus, Robert A.  
APPLICANT: Lee, Geoffrey F.  
TITLE OF INVENTION: No. 5736364e1 Factor VIIa Inhibitors  
NUMBER OF SEQUENCES: 58

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Genentech, Inc.  
;; STREET: 460 Point San Bruno Blvd  
;; CITY: South San Francisco  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 94080  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Winpacin (Genentech)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/566,800A  
;; FILING DATE: 04-Dec-1995  
;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Kubinec, Jeffrey S.  
;; REGISTRATION NUMBER: 36,575  
;; REFERENCE/DOCKET NUMBER: P0958B  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 415/225-8228  
;; TELEFAX: 415/952-9881  
;; TELEX: 910/371-7168  
;; INFORMATION FOR SEQ ID NO: 58:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 5 amino acids  
;; TYPE: Amino Acid  
;; TOPOLOGY: Linear  
;;  
;; US-08-566-800A-58

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5  
Db 1 GGGGS 5

;; RESULT 8  
;; US-08-244-469-5  
;; Sequence 5, Application US/08244469  
;; Patent No. 5736387  
;;  
;; GENERAL INFORMATION:  
;; APPLICANT: Paul, Ralph W.  
;; APPLICANT: Overell, Robert  
;; TITLE OF INVENTION: ENVELOPE FUSION VECTORS FOR USE IN GENE  
;; TITLE OF INVENTION: DELIVERY  
;; NUMBER OF SEQUENCES: 7  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: MORRISON & FOERSTER  
;; STREET: 755 PAGE MILL ROAD  
;; CITY: PALO ALTO  
;; STATE: CA  
;; COUNTRY: USA  
;; ZIP: 94304-1018  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/244,469  
;; FILING DATE: 01-JUN-1994  
;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Dylan, Tyler M.  
;; REGISTRATION NUMBER: 37,612  
;; REFERENCE/DOCKET NUMBER: 22627-20007.20  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 813-5600  
;; TELEFAX: (415) 494-0792

;; TELEX: 706141 MRSNPOERS SFO  
;; INFORMATION FOR SEQ ID NO: 5:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 5 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;;  
;; US-08-244-469-5

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5  
Db 1 GGGGS 5

;; RESULT 9  
;; US-08-189-331-140  
;; Sequence 140, Application US/08189331  
;; Patent No. 5747334  
;;  
;; GENERAL INFORMATION:  
;; APPLICANT: Kay, B. K.  
;; APPLICANT: Fowlkes, D. M.  
;; TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
;; NUMBER OF SEQUENCES: 186  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Pennie & Edmonds  
;; STREET: 1155 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: U.S.A.  
;; ZIP: 10036-2711  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/189,331  
;; FILING DATE: Concurrently herewith  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Mierock, S. Leslie  
;; REGISTRATION NUMBER: 18,972  
;; REFERENCE/DOCKET NUMBER: 1101-155  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212 790-9090  
;; TELEFAX: 212 869-8664/9741  
;; TELEX: 66141 PENNIE  
;; INFORMATION FOR SEQ ID NO: 140:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 5 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: peptide  
;;  
;; US-08-189-331-140

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5  
Db 1 GGGGS 5

;; RESULT 10  
;; US-08-333-577-6  
;; Sequence 6, Application US/08333577  
;; Patent No. 5786206

GENERAL INFORMATION:  
APPLICANT: Smith, Richard K.  
APPLICANT: Koduri, Raju  
APPLICANT: Young, Stephen G.  
APPLICANT: Wiltum, Joseph L.  
APPLICANT: Curtiss, Linda K.  
TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Goldsmith, Shore, Sutter &  
ADDRESSEE: Milamow, Ltd.  
STREET: 180 No. 5786206th Stetson, Suite 4700  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/333,577  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Gamson, Edward P.  
REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER: SCRP 234.0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312)616-5400  
TELEFAX: (312)616-5460  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-333-577-6

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
|||||  
Db 1 GGGGS 5

RESULT 11  
US-08-575-361A-32  
Sequence 32, Application US/08575361A  
Patent No. 5792640  
GENERAL INFORMATION:  
APPLICANT: Chandrasegaran, Srinivasan  
TITLE OF INVENTION: A GENERAL METHOD TO CLONE HYBRID  
TITLE OF INVENTION: RESTRICTION ENDONUCLEASES USING 11g GENE  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cushman Darby & Cushman L.L.P.  
STREET: 1100 New York Avenue, NW, Ninth Floor, East  
STREET: Tower  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/575,361A  
FILING DATE: 20-DEC-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kokulis, Paul N.  
REGISTRATION NUMBER: 16,773  
REFERENCE/DOCKET NUMBER: PNK/4130/213779/DJP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3000  
TELEFAX: 202-822-0944  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-575-361A-32

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
|||||  
Db 1 GGGGS 5

RESULT 12  
US-08-564-955-64  
Sequence 64, Application US/08564955  
Patent No. 5811238  
GENERAL INFORMATION:  
APPLICANT: STEMMER, WILLEM P.C.  
APPLICANT: CRAMERI, ANDREAS M.  
TITLE OF INVENTION: METHODS FOR GENERATING POLYNUCLEOTIDES  
TITLE OF INVENTION: HAVING DESIRED CHARACTERISTICS BY ITERATIVE SELECTION AND  
TITLE OF INVENTION: RECOMINATION  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW  
STREET: TWO EMBARCADERO CENTER, 8TH FLOOR  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: U.S.A.  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/564,955  
FILING DATE: 30-NOV-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/198,431  
FILING DATE: 17-FEB-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/537,874  
FILING DATE: 30-OCT-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/02126  
FILING DATE: 17-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: DUNN, TRACY J.  
REGISTRATION NUMBER: 34,587  
REFERENCE/DOCKET NUMBER: 165261-014611US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 64:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-564-955-64

Query Match 100.0%; Score 28; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
Db 1 GGGGS 5

RESULT 13  
US-08-528-523-13  
Sequence 13, Application US/08528523  
Patent No. 5824782  
GENERAL INFORMATION:  
APPLICANT: Hoeizer, Wolfgang  
APPLICANT: von Hoegen, Ilka  
APPLICANT: Strittmatter, Wolfgang  
APPLICANT: Matzku, Siegfried  
TITLE OF INVENTION: Immunconjugates II  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Millen, White, Zelano & Branigan, P.C.  
STREET: 2200 Clarendon Boulevard, Suite 1400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/528,523  
FILING DATE: 06-NOV-1992  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 94114572.4  
FILING DATE: 16-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Hamlet-King, Diana  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: Merck 1717  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
FRAGMENT TYPE: internal  
US-08-528-523-13

Query Match 100.0%; Score 28; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
Db 1 GGGGS 5

RESULT 14

US-08-537-874-62  
Sequence 62, Application US/08537874  
Patent No. 5830721  
GENERAL INFORMATION:  
APPLICANT: Stemmer, Willem P.C.  
APPLICANT: Cramer, Andreas  
TITLE OF INVENTION: DNA Mutagenesis by Random Fragmentation  
TITLE OF INVENTION: and Reassembly  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/537,874  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP PCT/US95/02126  
FILING DATE: 17-FEB-1995  
APPLICATION NUMBER: US 08/198,431  
FILING DATE: 17-FEB-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Liebeschuetz, Joe  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 018097-014610  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 62:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
US-08-537-874-62

Query Match 100.0%; Score 28; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
Db 1 GGGGS 5

RESULT 15  
US-08-448-418-86  
Sequence 86, Application US/08448418  
Patent No. 5837242  
GENERAL INFORMATION:  
APPLICANT: Holliger, Kaapar-Phillipp  
APPLICANT: Griffiths, Andrew D  
APPLICANT: Hoogenboom, Hendricus RJM  
APPLICANT: Malmqvist, Magnus  
APPLICANT: Marks, James D  
APPLICANT: McGuinness, Brian T  
APPLICANT: Pope, Anthony R  
APPLICANT: Prospero, Terence D  
APPLICANT: Winter, Gregory P  
TITLE OF INVENTION: Multivalent and Multispecific Binding  
TITLE OF INVENTION: Proteins, Their Manufacture and Use  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall O'Toole Gerstein Murray and Borun



STREET: 6300 Sears Tower 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/448,418  
FILING DATE: 14-MAY-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: C12N 15/62, 15/70, C07K 1/00  
APPLICATION NUMBER: PCT/GB93/02492  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9225453.1  
FILING DATE: 04-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9300816.7  
FILING DATE: 16-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 93303614.7  
FILING DATE: 10-MAY-1993  
PRIOR APPLICATION DATA: GB 9319969.3  
APPLICATION NUMBER: GB 9319969.3  
FILING DATE: 22-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: David W. Clough  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 28111/32651  
INFORMATION FOR SEQ ID NO: 86:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide linker  
US-08-448-418-86

Query Match 100.0%; Score 28; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGS 5  
|||||  
Db 1 GGGS 5

Search completed: November 21, 2004, 13:44:43  
Job time: 1.64781 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:42:32 ; Search time 2.18978 Seconds  
(without alignments)  
808.591 Million cell updates/sec

Title: US-09-936-702-1  
Perfect score: 28  
Sequence: 1 GCGGS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1570615 seqs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/ECTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	100.0	5	9	US-09-832-297A-1 Sequence 1, Appli
2	28	100.0	5	9	US-09-832-297A-2 Sequence 2, Appli
3	28	100.0	5	9	US-09-287-849-44 Sequence 44, Appli
4	28	100.0	5	9	US-09-147-142-31 Sequence 31, Appli
5	28	100.0	5	9	US-09-214-645-1 Sequence 1, Appli
6	28	100.0	5	9	US-09-858-616-2 Sequence 2, Appli
7	28	100.0	5	9	US-09-779-233-45 Sequence 45, Appli
8	28	100.0	5	9	US-09-989-789-3 Sequence 3, Appli
9	28	100.0	5	9	US-09-976-787-21 Sequence 21, Appli
10	28	100.0	5	9	US-09-192-854-180 Sequence 180, App
11	28	100.0	5	9	US-09-761-962-36 Sequence 36, Appli
12	28	100.0	5	9	US-09-333-527-5 Sequence 5, Appli
13	28	100.0	5	9	US-09-925-796-8 Sequence 8, Appli

14	28	100.0	5	9	US-09-815-837-116 Sequence 116, App
15	28	100.0	5	9	US-09-033-525-5 Sequence 5, Appli
16	28	100.0	5	9	US-09-779-451-7 Sequence 7, Appli
17	28	100.0	5	9	US-09-941-450-8 Sequence 8, Appli
18	28	100.0	5	9	US-09-818-247-25 Sequence 25, Appli
19	28	100.0	5	9	US-09-883-777-10 Sequence 10, Appli
20	28	100.0	5	9	US-09-867-262-3 Sequence 3, Appli
21	28	100.0	5	9	US-09-780-933-22 Sequence 22, Appli
22	28	100.0	5	9	US-09-480-238-10 Sequence 10, Appli
23	28	100.0	5	9	US-09-731-558-6 Sequence 6, Appli
24	28	100.0	5	9	US-09-828-708-123 Sequence 123, App
25	28	100.0	5	9	US-09-885-551A-3 Sequence 3, Appli
26	28	100.0	5	9	US-09-756-283A-14 Sequence 14, Appli
27	28	100.0	5	9	US-09-144-886-4 Sequence 4, Appli
28	28	100.0	5	9	US-09-999-745-56 Sequence 56, Appli
29	28	100.0	5	9	US-09-942-087A-8 Sequence 8, Appli
30	28	100.0	5	9	US-09-942-090-8 Sequence 8, Appli
31	28	100.0	5	9	US-09-554-000-40 Sequence 40, Appli
32	28	100.0	5	9	US-09-792-793A-1 Sequence 1, Appli
33	28	100.0	5	9	US-09-792-793A-2 Sequence 2, Appli
34	28	100.0	5	10	US-09-846-033B-212 Sequence 212, App
35	28	100.0	5	10	US-09-990-186-3 Sequence 3, Appli
36	28	100.0	5	10	US-09-897-844-8 Sequence 8, Appli
37	28	100.0	5	10	US-09-989-994-3 Sequence 3, Appli
38	28	100.0	5	10	US-09-911-261A-23 Sequence 23, Appli
39	28	100.0	5	10	US-09-942-024-84 Sequence 84, Appli
40	28	100.0	5	10	US-09-942-098-84 Sequence 84, Appli
41	28	100.0	5	10	US-09-969-748C-38 Sequence 38, Appli
42	28	100.0	5	10	US-09-992-124A-61 Sequence 61, Appli
43	28	100.0	5	10	US-09-949-039-37 Sequence 37, Appli
44	28	100.0	5	13	US-10-087-426-3 Sequence 3, Appli
45	28	100.0	5	13	US-10-057-505-15 Sequence 15, Appli

## ALIGNMENTS

RESULT 1  
US-09-832-297A-1  
; Sequence 1, Application US/09832297A  
; Publication No. US2001005566A1  
; GENERAL INFORMATION:  
; APPLICANT: Luitken, George A.  
; TITLE OF INVENTION: METHOD FOR VIEWING TUMOR TISSUE LOCATED WITHIN A BODY CAVITY  
; FILE REFERENCE: FLTOR1120-2  
; CURRENT APPLICATION NUMBER: US/09/832,297A  
; CURRENT FILING DATE: 2001-04-09  
; PRIOR APPLICATION NUMBER: US 09/362,805  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: US 09/173,190  
; PRIOR FILING DATE: 1998-10-15  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide linker moiety  
US-09-832-297A-1

Query Match 100.0%; Score 28; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 1 GCGGS 5  
|||  
DB 1 GCGGS 5

RESULT 2  
US-09-832-297A-2

```
; Sequence 2, Application US/09832297A
; Publication No. US20010055566A1
; GENERAL INFORMATION:
; APPLICANT: Fluoroprobe, Inc.
; APPLICANT: LUTKEN, George A.
; TITLE OF INVENTION: METHOD FOR VIEWING TUMOR TISSUE LOCATED WITHIN A BODY CAVITY
; FILE REFERENCE: FLUOR1120-2
; CURRENT APPLICATION NUMBER: US/09/832,297A
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 09/362,805
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: US 09/173,190
; PRIOR FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Peptide linker moiety. The sequence can be repeated n number of
; US-09-832-297A-2
```

```
Query Match          100.0%; Score 28; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGGGS 5
        |||||
Db       1 GGGGS 5
```

```
RESULT 3
US-09-287-849-44
; Sequence 44, Application US/09287849
; Patent No. US2002009459A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:flexible
; OTHER INFORMATION: polylinker
; US-09-287-849-44
```

```
Query Match          100.0%; Score 28; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGGGS 5
        |||||
Db       1 GGGGS 5
```

```
RESULT 4
US-09-147-142-31
; Sequence 31, Application US/09147142
; Patent No. US20020018749A1
; GENERAL INFORMATION:
; APPLICANT: HUDSON, Peter John
; APPLICANT: KORTT, Alex Andrew
; APPLICANT: IRVING, Robert Alexander
; APPLICANT: ATWELL, John Leslie
; TITLE OF INVENTION: HIGH AVIDITY POLYVALENT AND POLYSPECIFIC REAGENTS
; FILE REFERENCE: 016786/0212
; CURRENT APPLICATION NUMBER: US/09/147,142
; CURRENT FILING DATE: 1999-03-05
; EARLIER APPLICATION NUMBER: PCT/AU98/00212
; EARLIER FILING DATE: 1998-03-26
; EARLIER APPLICATION NUMBER: AU PO 5917
; EARLIER FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide linker
; US-09-147-142-31
```

```
Query Match          100.0%; Score 28; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGGGS 5
        |||||
Db       1 GGGGS 5
```

```
RESULT 5
US-09-214-645-1
; Sequence 1, Application US/09214645
; Patent No. US20020028443A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: METHOD OF DNA SHUFFLING WITH
; TITLE OF INVENTION: POLYNUCLEOTIDES PRODUCED BY BLOCKING OR INTERRUPTING A
; TITLE OF INVENTION: SYNTHESIS OR AMPLIFICATION PROCESS
; FILE REFERENCE: DIVER1220-2 US/09/214,645
; CURRENT APPLICATION NUMBER: US/09/214,645
; CURRENT FILING DATE: 1999-09-27
; PRIOR APPLICATION NUMBER: PCT/US97/12239
; PRIOR FILING DATE: 1997-07-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: linker peptide
; US-09-214-645-1
```

```
Query Match          100.0%; Score 28; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGGGS 5
        |||||
Db       1 GGGGS 5
```

RESULT 6  
US-09-858-616-2  
; Sequence 2, Application US/09858616  
; Patent No. US20020031771A1  
; GENERAL INFORMATION:  
; APPLICANT: DIVERSA CORPORATION  
; APPLICANT: SHORT, Jay  
; TITLE OF INVENTION: SEQUENCE BASED SCREENING  
; FILE REFERENCE: DIVER1210-6  
; CURRENT APPLICATION NUMBER: US/09/858,616  
; PRIOR FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: US 09/571,499  
; PRIOR FILING DATE: 2000-05-15  
; PRIOR APPLICATION NUMBER: US 09/557,276  
; PRIOR FILING DATE: 2000-04-24  
; PRIOR APPLICATION NUMBER: US 08/592,002  
; PRIOR FILING DATE: 1996-08-02  
; PRIOR APPLICATION NUMBER: US 60/008,317  
; PRIOR FILING DATE: 1995-12-07  
; PRIOR APPLICATION NUMBER: US 08/944,795  
; PRIOR FILING DATE: 1997-10-06  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Linker peptide  
US-09-858-616-2

Query Match 100.0%; Score 28; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGGGS 5  
DB 1 GGGGS 5  
RESULT 7  
US-09-779-233-45  
; Sequence 45, Application US/09779233  
; Patent No. US20020045158A1  
; GENERAL INFORMATION:  
; APPLICANT: Casey, Casey  
; TITLE OF INVENTION: CELLS FOR DRUG DISCOVERY  
; FILE REFERENCE: 8325-0010  
; CURRENT APPLICATION NUMBER: US/09/779,233  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 45  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: linker  
US-09-779-233-45

Query Match 100.0%; Score 28; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGGGS 5  
DB 1 GGGGS 5  
RESULT 8  
US-09-989-789-3

; Sequence 3, Application US/09989789  
; Patent No. US20020063379A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Qiang  
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE  
; FILE REFERENCE: 8325-0011.20 / 511-US2  
; CURRENT APPLICATION NUMBER: US/09/989,789  
; CURRENT FILING DATE: 2002-03-25  
; NUMBER OF SEQ ID NOS: 4085  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: peptide linker  
US-09-989-789-3

Query Match 100.0%; Score 28; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGGGS 5  
DB 1 GGGGS 5

RESULT 9  
US-09-976-787-21  
; Sequence 21, Application US/09976787  
; Patent No. US20020064528A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhu, Zhenping  
; APPLICANT: Witte, Larry  
; TITLE OF INVENTION: Antibodies Specific to KDR and Uses Thereof  
; FILE REFERENCE: 11245/46505  
; CURRENT APPLICATION NUMBER: US/09/976,787  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 09/493,539  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: US 60/117,726  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: WordPerfect 8.0 for Windows  
; SEQ ID NO 21  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: peptide linker  
US-09-976-787-21

Query Match 100.0%; Score 28; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGGGS 5  
DB 1 GGGGS 5

RESULT 10  
US-09-192-854-180  
; Sequence 180, Application US/09192854  
; Patent No. US20020068276A1  
; GENERAL INFORMATION:  
; APPLICANT: Winter, Greg  
; APPLICANT: Tomlinson, Ian  
; TITLE OF INVENTION: Methods for Selecting Functional Peptides  
; FILE REFERENCE: 3789/72916  
; CURRENT APPLICATION NUMBER: US/09/192,854  
; CURRENT FILING DATE: 1998-11-17

```
; EARLIER APPLICATION NUMBER: 60/066,729
; EARLIER FILING DATE: 1997-11-21
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 180
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Linker peptide
; OTHER INFORMATION: for connecting variable domains.
US-09-192-854-180

Query Match
Best Local Similarity 100.0%; Score 28; DB 9; Length 5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 11
US-09-761-962-36
; Sequence 36, Application US/09761962
; Patent No. US20020077285A1
; GENERAL INFORMATION:
; APPLICANT: Memorial Sloan-Kettering Cancer Center
; TITLE OF INVENTION: Identification and Characterization of Multiple Splice
; TITLE OF INVENTION: Variants of Mu-
; FILE REFERENCE: 830002-2000.1
; CURRENT APPLICATION NUMBER: US/09/761,962
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/743,872
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 36
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: basic unit of a linking peptide
US-09-761-962-36

Query Match
Best Local Similarity 100.0%; Score 28; DB 9; Length 5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 12
US-09-333-527-5
; Sequence 5, Application US/09333527
; Patent No. US20020078472A1
; GENERAL INFORMATION:
; APPLICANT: Paul CHRISTOU; Eva STROGER; Rainer FISCHER; Carmen MARTIN-VAQUERO; Stef
; TITLE OF INVENTION: METHODS AND MEANS FOR EXPRESSION OF MAMMALIAN POLYPEPTIDES
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.25 inch, 1.44mb
; COMPUTER: IBM PS/2
```

```
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/333,527
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/089,322
; FILING DATE: June 15, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: KL/JIC 202.1 - JEL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5
; TYPE: amino acid
; TOPOLOGY: linear
US-09-333-527-5

Query Match
Best Local Similarity 100.0%; Score 28; DB 9; Length 5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 13
US-09-925-796-8
; Sequence 8, Application US/09925796
; Patent No. US20020081614A1
; GENERAL INFORMATION:
; APPLICANT: Case, Casey C.
; APPLICANT: Zhang, Lei
; TITLE OF INVENTION: Functional Genomics Using Zinc Finger Proteins
; FILE REFERENCE: 019456-002000US
; CURRENT APPLICATION NUMBER: US/09/925,796
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/395,448
; PRIOR FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: 09/229,037
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:linker
US-09-925-796-8

Query Match
Best Local Similarity 100.0%; Score 28; DB 9; Length 5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
Db 1 GGGGS 5

RESULT 14
US-09-815-837-116
; Sequence 116, Application US/09815837
; Patent No. US20020082411A1
; GENERAL INFORMATION:
; APPLICANT: Carter, Darrick
```

```

; APPLICANT: Zhu, Shirley
; APPLICANT: Arimilli, Subhashini
; APPLICANT: Wang, Aijun
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Immune Mediators and Related Methods
; FILE REFERENCE: 014058-005670US
; CURRENT APPLICATION NUMBER: US/09/815,837
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,274
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 60/204,249
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: US 60/264,003
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 116
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:downstream
; OTHER INFORMATION: Linker for COS96
US-09-815-837-116

```

```

Query Match          100.0%; Score 28; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 GGGGS 5
        |||||
Db      1 GGGGS 5

```

```

RESULT 15
US-09-033-525-5
; Sequence 5, Application US/09033525
; Patent No. US20020090374A1
; GENERAL INFORMATION:
; APPLICANT: Yarkoni, Shai
; APPLICANT: Ben-Yehudah, Ahmi
; APPLICANT: Azar, Yehudith
; APPLICANT: Agellon, Rami
; APPLICANT: Belostotsky, Ruth
; APPLICANT: Loterdoun-Galski, Haya
; TITLE OF INVENTION: CHIMERIC PROTEINS WITH CELL-TARGETING
; FILE REFERENCE: 9457-009-999
; CURRENT APPLICATION NUMBER: US/09/033,525
; CURRENT FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Flexible polylinker
US-09-033-525-5

```

```

Query Match          100.0%; Score 28; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 GGGGS 5
        |||||
Db      1 GGGGS 5

```

```

Search completed: November 21, 2004, 14:03:48
Job time : 3.18978 secs

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:29:16 ; Search time 0.547445 Seconds  
(without alignments)  
878.780 Million cell updates/sec

Title: US-09-936-702-1  
Perfect score: 28  
Sequence: 1 GGGGS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	100.0	37	2 S29113	diptericin homolog
2	28	100.0	64	2 A86333	hypothetical prote
3	28	100.0	66	2 H84489	hypothetical prote
4	28	100.0	69	1 MIEC77	microcin B17 precu
5	28	100.0	78	2 E84686	hypothetical prote
6	28	100.0	80	2 T10550	hypothetical prote
7	28	100.0	81	2 PC2047	grain-sofiness pro
8	28	100.0	82	2 S19774	glycine-rich prote
9	28	100.0	85	2 T32664	hypothetical prote
10	28	100.0	92	2 P00743	grain-sofiness pro
11	28	100.0	97	2 T48330	hypothetical prote
12	28	100.0	100	2 T49621	hypothetical prote
13	28	100.0	102	2 T23332	hypothetical prote
14	28	100.0	104	2 T02612	hypothetical prote
15	28	100.0	104	2 JCA190	holotricin 3 precu
16	28	100.0	108	2 G86252	hypothetical prote
17	28	100.0	109	2 S58673	RNA-binding protei
18	28	100.0	110	2 AC2391	RNA-binding protei
19	28	100.0	114	2 S28821	transcription fact
20	28	100.0	115	2 T35387	hypothetical prote
21	28	100.0	119	2 T07695	hypothetical prote
22	28	100.0	120	2 A81109	hypothetical prote
23	28	100.0	120	2 D83415	hypothetical prote
24	28	100.0	122	2 T04118	hypothetical prote
25	28	100.0	122	2 D86754	mitochondrial proc
26	28	100.0	125	2 T16247	prophage p12 prote
27	28	100.0	128	2 T30428	hypothetical prote
28	28	100.0	131	2 H69062	molybdenum transpo
29	28	100.0	133	2 G75432	hypothetical prote

30	28	100.0	135	2 S55647	hypothetical prote
31	28	100.0	136	2 T02870	globulin 2 precurs
32	28	100.0	136	2 T23282	hypothetical prote
33	28	100.0	139	2 C87544	hypothetical prote
34	28	100.0	140	2 AC3088	hypothetical prote
35	28	100.0	144	2 S35716	glycine-rich prote
36	28	100.0	144	2 S04069	glycine-rich prote
37	28	100.0	144	2 T34730	probable gas vesic
38	28	100.0	145	1 UQ1062	glycine-rich prote
39	28	100.0	145	2 E84469	probable glycine-r
40	28	100.0	148	2 S46514	puroindoline-b pre
41	28	100.0	148	2 I38891	caudal-type homeot
42	28	100.0	149	2 T23179	hypothetical prote
43	28	100.0	150	2 C86224	hypothetical prote
44	28	100.0	152	2 T04811	stfII protein homo
45	28	100.0	155	2 C86206	hypothetical prote

## ALIGNMENTS

RESULT 1  
S29113  
diptericin homolog - flesh fly (Sarcophaga peregrina)  
C.Species: Sarcophaga peregrina  
C.Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C.Accession: S29113  
R.Ishikawa, M.; Kubo, T.; Natori, S.  
Biochem. J. 287, 573-578, 1992  
A.Title: Purification and characterization of a diptericin homologue from Sarcophaga per  
A.Reference number: S29113; MUID:93074996; PMID:1445217  
A.Accession: S29113  
A.Status: preliminary  
A.Molecule type: protein  
A.Residues: 1-37 <ISH>  
A.Cross-references: UNIPROT:Q9TWM2

Query Match 100.0%; Score 28; DB 2; Length 37;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
Db 19 GGGGS 23

## RESULT 2

A86333  
hypothetical protein T20H2.25 - Arabidopsis thaliana  
C.Species: Arabidopsis thaliana (mouse-ear cress)  
C.Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C.Accession: A86333  
R.Theologis, A.; Becker, J.R.; Palm, C.J.; Federgraff, N.A.; Kaul, S.; White, O.; Alonso, J.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, K.; Jensen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A.Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luvois, J.S.; Maiti, R.; Marziani, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A.Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A.Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A.Reference number: A86141; MUID:21016719; PMID:11130712  
A.Accession: A86333  
A.Molecule type: DNA  
A.Status: preliminary  
A.Residues: 1-64 <STO>  
A.Cross-references: UNIPROT:O9LNS5; GB:A8005172; NID:98779001; PIDN:AAF79916.1; GSPDB:GN  
C.Genetics:  
A.Map position: 1

Query Match 100.0%; Score 28; DB 2; Length 64;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
|||||  
Db 25 GGGGS 29

## RESULT 3

H84489  
hypothetical protein At2g10020 (imported) - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C/Accession: H84489

R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Umayam, L.; Tallon, L.

euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: H84489

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-66 <STO>

A/Cross-references: UNIPROT:Q9SK19; GB:AE002093; NID:g4556680; PIDN:AMD2697.1; GSPDB:GN

C/Genetics:

A/Gene: At2g10020

A/Map position: 2

Query Match 100.0%; Score 28; DB 2; Length 66;  
Best Local Similarity 100.0%; Pred. No. 2,2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
|||||  
Db 22 GGGGS 26

## RESULT 4

MIEC77

microcin B17 precursor - Escherichia coli plasmid pmcCB17

C/Species: Escherichia coli

C/Date: 30-Jun-1988 #sequence\_revision 01-Dec-1995 #text\_change 09-Jul-2004

C/Accession: A25219; A32058; I41099; A58368; S67977

R/Davagnino, J.; Herrero, M.; Furlong, D.; Moreno, F.; Kolter, R.

Proteins 1, 230-238, 1986

A/Title: The DNA replication inhibitor microcin B17 is a forty-three-amino-acid protein

A/Reference number: A25219; MUID:88217867; PMID:3329725

A/Accession: A25219

A/Molecule type: DNA

A/Residues: 1-69 <DAV>

A/Cross-references: UNIPROT:P05834; GB:M15469; NID:g146787; PIDN:AAA24141.1; PID:g146786

J./Genlind, O.; Moreno, F.; Kolter, R.

A/Title: DNA sequence, products, and transcriptional pattern of the genes involved in pT

A/Reference number: A32058; MUID:89123111; PMID:2644225

A/Accession: A32058

A/Molecule type: DNA

A/Residues: 1-69 <GEN>

A/Cross-references: GB:M24253; NID:g341145; PIDN:AAA72741.1; PID:g522250

R./Connell, N.; Han, Z.; Moreno, F.; Kolter, R.

Mol. Microbiol. 1, 195-201, 1987

A/Title: An E. coli promoter induced by the cessation of growth.

A/Reference number: I41099; MUID:88216163; PMID:2835580

A/Accession: I41099

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-14 <CON>

A/Cross-references: EMBL:X06417; NID:g41978; PIDN:CAA29725.1; PID:g41979

R./Li, Y.M.; Milne, J.C.; Madison, L.L.; Kolter, R.; Walsh, C.T.

Science 274, 1188-1193, 1996

A/Title: From peptide precursors to oxazole and thiazole-containing peptide antibiotics;

A/Reference number: A58368; MUID:97053605; PMID:8895467

A/Accession: A58368

A/Molecule type: protein

A/Residues: 27-38 <LTY>

A/Experimental source: Escherichia coli strain ZK4 (PY113)

A/Note: mass spectroscopy of peptides and biosynthetic intermediates

R./Torgay, P.; Lee, U.; Koedel, U.; Vivas, E.; Warner, P.; Jeebaratnam, D.; Kolter, R.

Proc. Natl. Acad. Sci. U.S.A. 91, 4519-4523, 1994

A/Title: Posttranslational modifications in microcin B17 define an additional class of DR

A/Reference number: A58375; MUID:94240167; PMID:8183941

A/Contents: annotation; (1)H-NMR spectroscopy of modified peptides

R./Bayer, A.; Freund, S.; Jung, G.

Eur. J. Biochem. 234, 414-426, 1995

A/Title: Post-translational heterocyclic backbone modifications in the 43-peptide antibiotic

A/Reference number: S67977; MUID:96128168; PMID:8536683

A/Accession: S67977

A/Status: preliminary

A/Molecule type: protein

A/Residues: 27-38 <BAT>

C/Genetics:

A/Gene: mcba

A/Genome: pmcCB17

A/Description: inhibits DNA gyrase, stopping DNA replication

A/Note: active against a large number of gram-negative enteric bacteria

C/Superfamily: microcin

C/Keywords: antibiotic; DNA replication inhibitor; oxazole/thiazole ring

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-69/Product: microcin B17 #status experimental <MNT>

F:39-40/Cross-link: oxazole (Gly-Ser) #status experimental

F:40-41/Cross-link: thiazole (Ser-Cys) #status experimental

F:47-48/Cross-link: thiazole (Gly-Cys) #status experimental

F:50-51/Cross-link: thiazole (Gly-Cys) #status experimental

F:54-55/Cross-link: thiazole (Gly-Cys) #status experimental

F:55-56/Cross-link: oxazole (Cys-Ser) #status experimental

F:61-62/Cross-link: oxazole (Gly-Ser) #status experimental

F:64-65/Cross-link: oxazole (Gly-Ser) #status experimental

Query Match 100.0%; Score 28; DB 1; Length 69;  
Best Local Similarity 100.0%; Pred. No. 2,3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
|||||  
Db 36 GGGGS 40

## RESULT 5

E84686

hypothetical protein At2g28570 (imported) - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C/Accession: E84686

R./Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Umayam, L.; Tallon, L.

euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: E84686

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-78 <STO>

A/Cross-references: UNIPROT:Q9SK01; GB:AE002093; NID:g4510404; PIDN:AMD21491.1; GSPDB:GNK

C/Genetics:

A/Gene: At2g28570

A/Map position: 2

Query Match 100.0%; Score 28; DB 2; Length 78;  
Best Local Similarity 100.0%; Pred. No. 2,6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
|||||  
Db 53 GGGGS 57

RESULT 6  
T10550  
hypothetical protein T12G13.70 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C:Accession: T10550  
R:Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro  
submitted to the Protein Sequence Database, June 1999  
A:Reference number: Z16533  
A:Accession: T10550  
A:Molecule type: DNA  
A:Residues: 1-80 <BEV>  
A:Cross-references: UNIPROT:Q9SUF7; EMBL:AF080252; GSPDB:GN00062; ATSP:T12G13.70  
C:Genetics:  
A:Gene: ATSP:T12G13.70  
A:Map position: 4

Query Match  
Best Local Similarity 100.0%; Score 28; DB 2; Length 80;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
|||||  
Db 27 GGGGS 31

RESULT 7  
PC2047  
grain-softness protein - wheat (fragments)  
C:Species: Triticum aestivum (common wheat)  
C:Date: 05-Aug-1994 #sequence\_revision 05-Aug-1994 #text\_change 09-Jul-2004  
C:Accession: PC2047  
R:Jolly, C.J.; Rahman, S.; Kortt, A.A.; Higgins, T.J.V.  
Theor. Appl. Genet. 86, 589-597, 1993  
A:Title: Characterisation of the wheat Mr 15000 grain-softness protein and analysis of t  
A:Reference number: P00743  
A:Accession: PC2047  
A:Molecule type: protein  
A:Residues: 1-18;19-24;25-31;32-38;39-45;46-51;52-56;57-60;61-65;66-71;72-77;78-81 <JOL>  
A:Cross-references: UNIPROT:Q7M1T8  
C:Comment: This protein is the product of the Ha locus and thus be the major factor that  
C:Keywords: seed

Query Match  
Best Local Similarity 100.0%; Score 28; DB 2; Length 81;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
|||||  
Db 4 GGGGS 8

RESULT 8  
S19774  
glycine-rich protein - tomato (fragment)  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
C:Accession: S19774  
R:Parsons, B.L.  
submitted to the EMBL Data Library, May 1991  
A:Reference number: S19773  
A:Accession: S19774  
A:Molecule type: mRNA  
A:Residues: 1-82 <PAR>  
A:Cross-references: UNIPROT:Q04130; EMBL:X59883; NID:gi19321; P1DN:CAA42538.1; P1D:gi19322  
C:Superfamily: glycine-rich RNA-binding protein; ribonucleoprotein repeat homology

Query Match  
Best Local Similarity 100.0%; Score 28; DB 2; Length 82;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
|||||  
Db 72 GGGGS 76

RESULT 9  
T32664  
hypothetical protein F16B4.7 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T32664  
R:Davidson, S.; Wohlmann, P.; Bauer, C.; O'Neal, D.  
submitted to the EMBL Data Library, December 1997  
A:Description: The sequence of C. elegans cosmid F16B4.  
A:Reference number: Z21208  
A:Accession: T32664  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-85 <DAV>  
A:Cross-references: UNIPROT:Q44629; EMBL:AF039048; P1DN:AAE94238.1; GSPDB:GN00023; CESP:  
A:Experimental source: strain Bristol N2; clone F16B4  
C:Genetics:  
A:Gene: CESP:F16B4.7  
A:Map position: 5  
A:introns: 36/1  
C:Superfamily: Arabidopsis glycine-rich protein 3

Query Match  
Best Local Similarity 100.0%; Score 28; DB 2; Length 85;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
|||||  
Db 39 GGGGS 43

RESULT 10  
P00743  
grain-softness protein - wheat (fragments)  
C:Species: Triticum aestivum (common wheat)  
C:Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 09-Jul-2004  
C:Accession: P00743  
R:Jolly, C.J.; Rahman, S.; Kortt, A.A.; Higgins, T.J.V.  
Theor. Appl. Genet. 86, 589-597, 1993  
A:Title: Characterisation of the wheat Mr 15000 grain-softness protein and analysis of t  
A:Reference number: P00743  
A:Accession: P00743  
A:Molecule type: protein  
A:Residues: 1-92 <JOL>  
A:Cross-references: UNIPROT:Q7M1T7  
A:Experimental source: seed  
C:Keywords: seed

Query Match  
Best Local Similarity 100.0%; Score 28; DB 2; Length 92;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
|||||  
Db 4 GGGGS 8

RESULT 11  
T48330  
hypothetical protein F15A17.120 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C:Accession: T48330  
R:Bevan, M.; Terry, N.; Ardiles, W.; Buyschaert, C.; Dasseville, R.; De Clerck, R.; De  
ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.  
submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z24491

A/Accession: T48330  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-97 <BEV>  
A/Cross-references: UNIPROT:Q9LYX6; EMBL:AL163002  
A/Experimental source: cultivar Columbia; BAC clone F15A17  
C/Genetics:  
A/Map position: 5  
A/Intons: 7/1  
A/Note: F15A17.120

Query Match 100.0%; Score 28; DB 2; Length 97;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
Db 47 GGGGS 51

## RESULT 12

T49621  
hypothetical protein B5022.30 [imported] - Neurospora crassa  
C/Species: Neurospora crassa  
C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C/Accession: T49621  
R/Schulte, U.; Alyn, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, May 2000  
A/Reference number: Z25022  
A/Accession: T49621  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-100 <SCH>  
A/Cross-references: EMBL:AJ355932; GSPDB:GN00116; NCSP:B5022.30  
A/Experimental source: BAC clone B5022; strain OR74A  
C/Genetics:  
A/Gene: NCSP:B5022.30  
A/Map position: 6  
A/Intons: 22/1; 52/1

Query Match 100.0%; Score 28; DB 2; Length 100;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
Db 11 GGGGS 15

## RESULT 13

T25332  
hypothetical protein T26H5.4 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T25332  
R/Gardner, A.  
submitted to the EMBL Data Library, November 1996  
A/Reference number: Z20017  
A/Accession: T25332  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-102 <WIL>  
A/Cross-references: UNIPROT:O62385; EMBL:Z82056; PIDN:CA804855.1; GSPDB:GN00023; CESP:T2  
C/Experimental source: clone T26H5  
C/Genetics:  
A/Gene: CESP:T26H5.4  
A/Map position: 5  
A/Intons: 13/1; 96/1  
C/Superfamily: hypothetical protein K01D12.8

Query Match 100.0%; Score 28; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
Db 52 GGGGS 56

## RESULT 14

T02612  
hypothetical protein At2g26120 [imported] - Arabidopsis thaliana  
N/Alternate names: hypothetical protein T19L18.7  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C/Accession: T02612; F84656  
R/Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.  
submitted to the EMBL Data Library, August 1998  
A/Description: Arabidopsis thaliana chromosome II BAC T19L18 genomic sequence.  
A/Reference number: Z14681  
A/Accession: T02612  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-104 <ROU>  
A/Cross-references: UNIPROT:O80985; EMBL:AC004747; NID:G3413696; PID:G3413702  
A/Experimental source: cultivar Columbia  
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; B  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; MUID:20083487; PMID:10617197  
A/Accession: F84656  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-104 <STO>  
A/Cross-references: GB:AE002093; NID:G3413702; PIDN:AA031225.1; GSPDB:GN00139  
C/Genetics:  
A/Gene: T19L18.7; At2g26120  
A/Map position: 2  
A/Intons: 49/3

Query Match 100.0%; Score 28; DB 2; Length 104;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
Db 96 GGGGS 100

## RESULT 15

JC4190  
holotricin 3 precursor - Holotrichia diomphalia  
N/Alternate names: antifungal protein  
C/Species: Holotrichia diomphalia  
C/Date: 04-Oct-1995 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C/Accession: JC4190  
R/Lee, S.Y.; Moon, H.J.; Kurata, S.; Natori, S.; Lee, B.L.  
Biol. Pharm. Bull. 18, 1049-1052, 1995  
A/Title: Purification and cDNA cloning of an antifungal protein from the hemolymph of Hol  
A/Reference number: JC4190; MUID:96073722; PMID:8535353  
A/Accession: JC4190  
A/Molecule type: mRNA  
A/Residues: 1-104 <LEE>  
A/Cross-references: UNIPROT:Q25055; DDBJ:D13744; NID:G1088433; PIDN:BA02889.1; PID:d100;  
C/Comment: This protein is a Gly- and His-rich protein and a constitutive protein of lar  
C/Keywords: hemolymph  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-104/Product: holotricin 3 #status predicted <MAT>

Query Match 100.0%; Score 28; DB 2; Length 104;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5

|||||  
Db 69 GGGG 73

Search completed: November 21, 2004, 13:43:26  
Job time : 2.54745 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:28:31 ; Search time 2.77372 Seconds  
(without alignments)  
1037.189 Million cell updates/sec

Title: US-09-936-702-1  
Perfect score: 28  
Sequence: 1 GGGGS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues  
Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 02: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	100.0	10	06X2S9	06X2S9 homo sapien
2	28	100.0	10	AAP83315	AAP83315 homo sapi
3	28	100.0	17	PKX5_PRRAM	P82617 periplaneta
4	28	100.0	17	09R4Y9	09R4Y9 alcalligenes
5	28	100.0	17	064450	064450 mus musculu
6	28	100.0	20	08W238	08W238 glycinine max
7	28	100.0	20	09R582	09R582 neisseria m
8	28	100.0	23	09UC00	09UC00 homo sapien
9	28	100.0	26	042226	042226 arabidopsis
10	28	100.0	31	09P115	09P115 enterococcu
11	28	100.0	37	09TWM2	09TWM2 sarcophaga
12	28	100.0	40	013832	013832 homo sapien
13	28	100.0	40	013833	013833 homo sapien
14	28	100.0	44	06Y1D5	06Y1D5 lactuca sat
15	28	100.0	44	AAR22556	AAR22556 lactuca s
16	28	100.0	45	07XMY7	07XMY7 oryza sativ
17	28	100.0	48	07QSB6	07QSB6 giardia lam
18	28	100.0	50	084Y55	084Y55 oryza sativ
19	28	100.0	50	094319	094319 oryza sativ
20	28	100.0	50	098C82	098C82 rhizobium l
21	28	100.0	53	07S2G8	07S2G8 neuropeptid
22	28	100.0	53	08LNM3	08LNM3 oryza sativ
23	28	100.0	54	06YV44	06YV44 oryza sativ
24	28	100.0	54	BAD17456	BAD17456 oryza sat
25	28	100.0	55	0751X3	0751X3 oryza sativ
26	28	100.0	55	AAS07084	AAS07084 oryza sat
27	28	100.0	60	08H647	08H647 oryza sativ
28	28	100.0	61	082GM4	082GM4 streptomyce
29	28	100.0	62	08N222	08N222 homo sapien
30	28	100.0	62	07E2W2	07E2W2 oryza sativ
31	28	100.0	62	BAD01701	BAD01701 oryza sat

32	28	100.0	62	2	BAD03150	BAD03150 oryza sat
33	28	100.0	63	2	094E89	094E89 oryza sativ
34	28	100.0	63	2	08WB65	08WB65 white spot
35	28	100.0	64	2	06K665	06K665 oryza sativ
36	28	100.0	64	2	09LNS5	09LNS5 arabidopsis
37	28	100.0	64	2	06ERC8	06ERC8 oryza sativ
38	28	100.0	65	2	0625J7	0625J7 oryza sativ
39	28	100.0	65	2	094J57	094J57 oryza sativ
40	28	100.0	65	2	BAD17269	BAD17269 oryza sat
41	28	100.0	66	2	09XZQ8	09XZQ8 caenorhabdi
42	28	100.0	66	2	09XT04	09XT04 macropus ru
43	28	100.0	66	2	08H3Q4	08H3Q4 oryza sativ
44	28	100.0	66	2	09SK19	09SK19 arabidopsis
45	28	100.0	67	2	06YVW2	06YVW2 oryza sativ

## ALIGNMENTS

RESULT 1	06X2S9	PRELIMINARY;	PRT;	10 AA.
ID	06X2S9			
AC	06X2S9			
DT	05-JUL-2004 (TREMBLrel. 27, Created)			
DT	05-JUL-2004 (TREMBLrel. 27, Last sequence update)			
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)			
DE	Translocase of the inner mitochondrial membrane (Fragment).			
GN	Name=TIM23;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid:9606;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Skin;			
RX	MEDLINE=22890074; PubMed=14527731;			
RA	Meyer R.G., Meyer-Ficca M.L., Jacobson E.L., Jacobson M.K.,			
RT	"Human poly(ADP-ribose) glycohydrolase (PARC) gene and the common			
RT	translocase 23 (TIM23).";			
RL	Gene 314:181-190(2003).			
DR	EMBL; AY258588; AAP83315.1; -.			
FT	NON_TER			
SO	SEQUENCE	10 AA;	893 MW;	2D7A7165B8787878 CRC64;
Query Match		100.0%;	Score 28;	DB 2;
Best Local Similarity		100.0%;	Pred. No. 2.2e+02;	
Matches	5;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;
QY	1 GGGGS 5			
DB	3 GGGGS 7			
RESULT 2	AAP83315	PRELIMINARY;	PRT;	10 AA.
ID	AAP83315			
AC	AAP83315			
DT	02-MAR-2004 (TREMBLrel. 27, Created)			
DT	02-MAR-2004 (TREMBLrel. 27, Last sequence update)			
DT	02-MAR-2004 (TREMBLrel. 27, Last annotation update)			
DE	Translocase of the inner mitochondrial membrane (Fragment).			
GN	Tim23.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid:9606;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Skin;			
RA	Meyer R.G., Meyer-Ficca M.L., Jacobson E.L., Jacobson M.K.,			
RT	"Human poly(ADP-ribose) glycohydrolase (PARC) gene and the common			
RT	promoter sequence it shares with inner mitochondrial membrane			

RT translocase 23 (TIM23).";  
 RL Gene 314:181-190(2003).  
 DR EMBL: AY258568; AAP8335.1; -.  
 FT NON TER 10 10  
 SQ SEQUENCE 10 AA; 893 MW; 2D7A7165B8787878 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 Db 3 GGGGS 7

## RESULT 3

ID PKRS\_PERAM STANDARD; PRT; 17 AA.  
 AC P82617;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 05-JUN-2004 (Rel. 44, Last annotation update)  
 DE Pyrokinin-5 (Pea-PK-5) (FKPRL-amide).  
 OS Periplaneta americana (American cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygota;  
 CC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattellidae;  
 OC Blattellidae; Periplaneta.  
 NC NCBI\_TaxID=6978;

RM [1]  
 RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
 RC TISSUE=Abdominal perisymphathetic organs;  
 RX MEDLINE=99212469; PubMed=10196736;  
 RA Predel R., Kellner R., Nachman R.J., Holman G.M., Rapus J., Gaede G.;  
 RT "Differential distribution of pyrokinin-isoforms in cerebral and  
 abdominal neurohemal organs of the American cockroach.";  
 RL Insect Biochem. Mol. Biol. 29:139-144(1999).  
 RN [2]

RP TISSUE SPECIFICITY.  
 RX MEDLINE=20189894; PubMed=10723010;  
 RA Predel R., Eckert M.;  
 RT "Tissue-specific distribution of FKPRamide in the nervous system of  
 the American cockroach.";  
 RL J. Comp. Neurol. 419:352-363(2000).  
 CC -1- FUNCTION: Mediates visceral muscle contractile activity (myotropic  
 activity).  
 CC -1- TISSUE SPECIFICITY: Mainly in abdominal perisymphathetic organs and  
 to a lesser extent in retrocerebral complex.  
 CC -1- MASS SPECTROMETRY: MW=1651.7; METHOD=MALDI; RANGE=1-17;  
 CC NOTE=Ref.1.  
 CC -1- SIMILARITY: Belongs to the pyrokinin family.  
 DR InterPro: IPR001484; Pyrokinin.  
 DR PROSITE: PS00539; PYROKININ; 1.  
 KW Annotation: Direct protein sequencing; Neuropeptide; Pyrokinin.  
 FT MOD\_RSS 17 17  
 FT MOD\_RSS 17 17  
 SQ SEQUENCE 17 AA; 1653 MW; 8527162BA45BA54 CRC64;

Query Match 100.0%; Score 28; DB 1; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 Db 1 GGGGS 5

## RESULT 4

ID Q9R4Y9 PRELIMINARY; PRT; 17 AA.  
 AC Q9R4Y9;  
 DT 01-MAY-2000 (TRENBLREL. 13, Created)  
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
 DT 01-JUN-2000 (TRENBLREL. 14, Last annotation update)  
 DE Aromatic amine dehydrogenase beta subunit (Fragment).

OS Alcaligenes faecalis.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 CC Alcaligenaceae; Alcaligenes.  
 CX NCBI\_TaxID=511;  
 RN [1]

RP SEQUENCE.  
 RX MEDLINE=94245619; PubMed=8188594;  
 RA Govindaraj S., Eisenstein E., Jones L.H., Sanders-Loehr J.,  
 RA Chistoserdov A.Y., Davidson V.L., Edwards S.L.;  
 RT "Aromatic amine dehydrogenase, a second tryptophan tryptophylquinone  
 enzyme.";  
 RL J. Bacteriol. 176:2922-2929(1994).  
 SQ SEQUENCE 17 AA; 1510 MW; 6EEBAEB9D89D2661 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 Db 2 GGGGS 6

## RESULT 5

ID Q64450 PRELIMINARY; PRT; 17 AA.  
 AC Q64450;  
 DT 01-NOV-1996 (TRENBLREL. 01, Created)  
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)  
 DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)  
 DE Uridine kinase (EC 2.7.1.48) (Fragment).  
 GN Name=Umpk;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Ropp P.A., Traut T.W.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U57332; AAB01998.1; -.  
 DR MGI: 98904; Umpk.  
 DR GO: GO:0016301; P:kinase activity; IEA.  
 DR GO: GO:0016740; F:transferase activity; IEA.  
 DR GO: GO:0004849; F:uridine kinase activity; IEA.  
 KW Kinase; Transferase.  
 FT NON\_TER 17 17  
 SQ SEQUENCE 17 AA; 1464 MW; 14E427CBA1168634 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 Db 5 GGGGS 9

## RESULT 6

ID Q8W238 PRELIMINARY; PRT; 20 AA.  
 AC Q8W238;  
 DT 01-MAR-2002 (TRENBLREL. 20, Created)  
 DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)  
 DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)  
 DE GT-2 factor (Fragment).  
 OS Glycine max (Soybean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC euosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 CX NCBI\_TaxID=3847;  
 RN [1]



RP SEQUENCE FROM N.A.  
 RX MEDLINE=21471140; PubMed=11587508;  
 RA O'Grady K., Gokhjian V.H., Naim C.J., Nagao R.T., Key J.L.;  
 RT "The transcript abundance of GmGT-2, a new member of the GT-2 family  
 of transcription factors from soybean, is down-regulated by light in a  
 RT phytochrome-dependent manner.";  
 RL Plant Mol. Biol. 47:367-378(2001).  
 DR EMBL; AF372500; AAL65126.1; -.  
 FT NON TER 20 20  
 SQ SEQUENCE 20 AA; 1692 MW; F65C75CD9C6B63B CRC64;

Query Match 100.0%; Score 28; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 DB 11 GGGGS 15

RESULT 7  
 Q9RS82 PRELIMINARY; PRT; 20 AA.  
 AC Q9RS82;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Transferrin-binding protein 2 (Fragment).  
 OS Neisseria meningitidis.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
 OC Neisseriaceae; Neisseria.  
 OX NCBI\_TaxID=487;  
 RN [1]  
 RP SEQUENCE.  
 RA MEDLINE=93307625; PubMed=8319886;  
 RA Griffiths E., Stevenson P., Byfield P., Ala-Aldeen D.A.,  
 RA Borriello S.P., Holland J., Parsons T., Williams P.;  
 RT "Antigenic relationships of transferrin-binding proteins from  
 RT Neisseria meningitidis, N. gonorrhoeae and Haemophilus influenzae:  
 RT cross-reactivity of antibodies to NH2-terminal peptides.";  
 RL FEMS Microbiol. Lett. 109:85-91(1993).  
 SQ SEQUENCE 20 AA; 1977 MW; 6000EB169F09227E CRC64;

Query Match 100.0%; Score 28; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 DB 3 GGGGS 7

RESULT 8  
 Q9UC00 PRELIMINARY; PRT; 23 AA.  
 AC Q9UC00;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Enhancement of wound HEALING process.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95130623; PubMed=7829572;  
 RA Pierschbacher M.D., Polarek J.W., Craig W.S., Tschopp J.F.,  
 RA Sipes N.J., Harper J.R.;  
 RL J. Cell. Biochem. 56:150-154(1994).  
 DR GO; GO:0009611; P:response to wounding; TAS.  
 SQ SEQUENCE 23 AA; 2268 MW; CE73999CB9903891 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 DB 11 GGGGS 15

RESULT 9  
 Q42226 PRELIMINARY; PRT; 26 AA.  
 AC Q42226;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Seed maturation protein (Fragment).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Dry seeds of A. thaliana ecotype Columbia;  
 RA CNRS;  
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Dry seeds of A. thaliana ecotype Columbia;  
 RA Raynal M., Grellet F., Lande M., Meyer Y., Cooke R., Delsemy M.;  
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z29850; CAA82818.1; -.  
 FT NON TER 1 1  
 SQ SEQUENCE 26 AA; 2370 MW; 6E0902E39464466A CRC64;

Query Match 100.0%; Score 28; DB 2; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 DB 11 GGGGS 15

RESULT 10  
 Q9FI15 PRELIMINARY; PRT; 31 AA.  
 AC Q9FI15;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Hypothetical protein EP0010.  
 GN Name=EP0010;  
 OS Enterococcus faecalis (Streptococcus faecalis).  
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.  
 OX NCBI\_TaxID=1351;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20453452; PubMed=10998166;  
 RA De Boever E.H., Clewell D.B., Fraser C.M.;  
 RT "Enterococcus faecalis conjugative plasmid PM373: complete nucleotide  
 RT sequence and genetic analyses of sex pheromone response.";  
 RL Mol. Microbiol. 37:1327-1341(2000).  
 DR EMBL; AB002565; AAG40421.1; -.  
 KW Hypothetical protein; Plasmid.  
 SQ SEQUENCE 31 AA; 3509 MW; 4E19CB94B3DB9421 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 6.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 Db 18 GGGGS 22

## RESULT 11

Q9TWM2 PRELIMINARY; PRT; 37 AA.  
 ID 09TWM2  
 AC 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Dipterician homolig (Fragment).  
 OS Sarcophaga peregrina (Flesh fly) (Boettcherisara peregrina).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;  
 OC Sarcophagidae; Sarcophaga.  
 NCBI\_TaxID=7386;  
 OX NCBI\_TaxID=7386;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=LARVAL HEMOLYMPH;  
 RX MEDLINE=93074996; PubMed=1445217;  
 RA Ishikawa M., Kudo T., Natori S.;  
 RT "Purification and characterization of a dipterician homologue from  
 RT Sarcophaga peregrina (Flesh fly).";  
 RL Biochem. J. 287:573-578 (1992).  
 CC -1- FUNCTION: BACTERICIDAL ACTIVITY AGAINST GRAM-NEGATIVE BACTERIA  
 CC E. COLI AND S. SONNET.  
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED BY THE FAT BODY AND IS SECRETED  
 CC INTO THE HEMOLYMPH.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION IN THE LARVAE STARTS A FEW MINUTES  
 CC AFTER THE INJURY OF THE BODY WALL REACHING A MAXIMUM AFTER ABOUT  
 CC 10 HOURS. THE MAXIMUM LASTS FOR AT LEAST 3 HOURS.  
 CC -1- INDUCTION: IN RESPONSE TO INJURY OF THE BODY WALL OF THE LARVAE.  
 DR PIR; S29113; S29113.  
 DR GO; GO:0006952; P:defense response; IEA.  
 DR GO; GO:0006805; P:xenobiotic metabolism; IEA.  
 KM Antibiotic; Insect immunity.  
 FT DOMAIN 18 22 POLY-GLY.  
 FT NON\_TER 37 37  
 SQ SEQUENCE 37 AA; 3928 MW; E3BAC8105D2DABC7 CRC64;  
 Query Match 100.0%; Score 28; DB 2; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 7; 7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGGGS 5  
 |||||  
 Db 19 GGGGS 23

RESULT 12  
 Q13832 PRELIMINARY; PRT; 40 AA.  
 ID 013832  
 AC 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last annotation update)  
 DE B2 bradykinin receptor basal promoter, allele BP-58-C (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96209920; PubMed=8655154;  
 RA Braun A., Maier E., Kammerer S., Mueller B., Roscher A.A.;  
 RT "A novel sequence polymorphism in the promoter region of the human  
 RT bradykinin B2-receptor gene.";  
 RL Hum. Genet. 97:688-689 (1996).  
 DR EMBL; X91663; CA62851.1; -;  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 KM Receptor.

FT NON\_TER 1 1  
 FT NON\_TER 40 40  
 SQ SEQUENCE 40 AA; 4140 MW; 3908E9AD371EE17F CRC64;  
 Query Match 100.0%; Score 28; DB 2; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 8; 3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 Db 6 GGGGS 10

## RESULT 13

Q13833 PRELIMINARY; PRT; 40 AA.  
 ID 013833  
 AC 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE B2 bradykinin receptor basal promoter, allele BP-58-T (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96209920; PubMed=8655154;  
 RA Braun A., Maier E., Kammerer S., Mueller B., Roscher A.A.;  
 RT "A novel sequence polymorphism in the promoter region of the human  
 RT bradykinin B2-receptor gene.";  
 RL Hum. Genet. 97:688-689 (1996).  
 DR EMBL; X91664; CA62852.1; -;  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 KM Receptor.  
 FT NON\_TER 1 1  
 FT NON\_TER 40 40  
 SQ SEQUENCE 40 AA; 4153 MW; 1408E9AD371EE17F CRC64;  
 Query Match 100.0%; Score 28; DB 2; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 8; 3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGGGS 5  
 |||||  
 Db 6 GGGGS 10

RESULT 14  
 Q6Y1D5 PRELIMINARY; PRT; 44 AA.  
 ID 06Y1D5  
 AC 06Y1D5;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE 60S acidic ribosomal protein P3 (Fragment).  
 OS Lactuca sativa (Garden lettuce).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asteridae;  
 OC Campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae;  
 OC Lactuca.  
 NCBI\_TaxID=4236;  
 OX NCBI\_TaxID=4236;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX Kuang H., Nevo E., Michelmore R.W.;  
 RA Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY193418; AAR22556.1; -;  
 DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
 KM Ribosomal protein.  
 FT NON\_TER 1 1  
 FT NON\_TER 44 44  
 SQ SEQUENCE 44 AA; 4358 MW; A9FF3A8D93744CD9 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 44;  
 Best Local Similarity 100.0%; Pred. No. 9.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 Db 39 GGGGS 43

## RESULT 15

AA22556 PRELIMINARY; PRT; 44 AA.  
 ID AAR22556  
 AC AAR22556;  
 DT 02-MAR-2004 (TRENBLREL. 27, Created)  
 DT 02-MAR-2004 (TRENBLREL. 27, Last sequence update)  
 DT 02-MAR-2004 (TRENBLREL. 27, Last annotation update)  
 DS 60S acidic ribosomal protein P3 (Fragment).  
 OS Lactuca sativa (Garden lettuce).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae;  
 OC Lactuca.  
 OK NCBI\_taxID=4236;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Calmar;  
 RA Kuang H., Nevo E., Michelmore R.W.;  
 RT "Heterogeneous Rates of Evolution at the Major Cluster Disease  
 Resistance Genes in Lettuce due to Distinct Genetic Processes."  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY193418; AAR22556.1; -.  
 KW Ribosomal protein.  
 FT NON\_TER 1 1  
 FT 44 44  
 SQ SEQUENCE 44 AA; 4358 MW; A9F3A8D93744CD9 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 44;  
 Best Local Similarity 100.0%; Pred. No. 9.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 Db 39 GGGGS 43

Search completed: November 21, 2004, 13:42:21  
 Job time : 5.77372 secs

**THIS PAGE BLANK (USPTO)**







XX	Pain-relieving target construct, linker moiety #12.
DE	
XX	Pain relief; tumour; therapeutic; breast cancer; lung cancer;
KW	Hodgkin's lymphoma; neuroblastoma; ischaemic.
XX	
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 3..57 /note= "Each tetra-glycine repeat can optionally be di-
FT	glycine or tri-glycine"
FT	Misc-difference 8..57 /note= "Optionally absent"
FT	Misc-difference 58 /note= "Any amino acid"
FN	
PN	MO200147512-AZ.
PD	
XX	05-JUL-2001.
PF	
PR	06-DEC-2000; 2000WO-US042661.
XX	
PR	08-DEC-1999; 99US-00457498.
PA	(FLUO-) FLUORO PROBE INC.
XX	
PI	Lutken GA,
XX	
DR	WPI: 2001-465198/50.
PT	Treatment of pain associated with an interior disease site, involves
XX	administering a pain-relieving target construct to the patient.
PS	Disclosure; Page 21; 31pp; English.
XX	
CC	The sequence represents the amino acid sequence of linker moiety #12 used
CC	to make a pain-relieving target construct. The construct comprises a pain
CC	-relieving agent linked to a ligand moiety that selectively binds to or
CC	is taken up by the tissue associated with the painful interior disease
CC	site. The construct is used for treating pain associated with an interior
CC	disease site. Since the pain-relieving agent is delivered by the ligand
CC	to the disease site intractable pain situated in the interior of the body
CC	such as caused by various tumours, such as breast cancer, lung cancer,
CC	Hodgkin's and non-Hodgkin's lymphoma, and neuroblastoma, and ischaemic
CC	and diseased tissues can be managed using a lower level of the pain
CC	relieving agent than is required when the pain-relieving agent is
XX	injected in the free state
SO	Sequence 60 AA;
Query Match	100.0%; Score 196; DB 4; Length 60;
Best Local Similarity	100.0%; Pred. No. 1e-11; Mismatches 0; Indels 0; Gaps 0
Matches 35; Conservative 0;	
OY	1 GGCGSGGCGSGGCGSGGCGSGGCGSGGCGS 35 
Db	3 GGCGSGGCGSGGCGSGGCGSGGCGSGGCGS 37
RESULT 7	
ID	AAG77867
AC	AAG77867 standard; protein; 60 AA.
XX	
AA	AAG77867;
XX	
DT	08-MAY-2002 (first entry)
XX	
DE	MHC class II H domain spacer.
XX	
KM	MHC class II H domain spacer; IG3-GIHLA fusion protein;
KM	major histocompatibility complex; MHC-peptide complex; MHC;
KW	MHC class I alpha chain; beta-2 microglobulin; MHC class II alpha chain.

KM	MHC class II beta chain; vaccine; immune response modulation;
KM	hyperproliferative disorder; neoplasm; hypermaglobulinaemia;
KM	viral infection; hepatitis; meningitis; bacterial infection;
KM	tuberculosis; gingivitis; parasitic infection; autoimmune disease;
KM	Hashimoto's disease; Graves' disease; rheumatoid arthritis; allergy;
KM	asthma; organ rejection; graft-versus-host disease; GVHD; HLA;
KM	human leukocyte antigen.
XX	Synthetic.
OS	
XX	
PN	WO200178768-A2.
XX	
PD	25-OCT--2001.
XX	
PF	12-APR-2001; 2001WO-US011912.
XX	
PR	12-APR-2000; 2000US-0196472P.
XX	
PA	(UNRP ) UNIV ROCHESTER.
XX	
Pt	Zauderer M, Smith ES;
DR	WPI, 2001-602927/68.
XX	
PT	Novel compound comprising major histocompatibility complex-peptide complexes, used to modulate immune responses.
PS	Example 3; Page 146; 166pp; English.
XX	
CC	The present sequence represents a major histocompatibility complex (MHC) class II H domain spacer. This protein was used as a replacement for the H domain present in MHC class II molecules in an example of the invention detailing the construction of human IgG3-HLA fusion proteins. Absence of the H domain or a suitable spacer would result in misfolding of the MHC class II molecule and the correct MHC binding site would not be formed. The invention comprises a compound which contains one or more major histocompatibility complex (MHC)-peptide complexes, and an antibody specific for a cell surface marker. The complexes comprise an MHC class I alpha chain, a beta-2 microglobulin molecule and an antigenic peptide bound in the MHC groove. Alternatively, the complexes may comprise an MHC class II alpha chain, an MHC class II beta chain, and an antigenic peptide bound in the MHC groove. The complexes are linked to the carboxyl terminus of the antibody. The compounds of the invention can be used as a vaccine to modulate an immune response. The compounds of the invention are useful for treating hyperproliferative disorders (e.g. neoplasms and hypergammaglobulinaemia); viral infections (e.g. hepatitis and meningitis); bacterial infections (e.g. tuberculosis and gingivitis); parasitic infections; autoimmune diseases (e.g. Hashimoto's disease, Graves' disease and rheumatoid arthritis); allergic reactions/conditions (e.g. asthma). The compounds of the invention may also be used in the treatment of organ rejection or graft-versus-host disease (GVHD)
BQ	Sequence 60 AA:
OY	Query Match 100.0%; Score 196; DB 4; Length 60; Best Local Similarity 100.0%; Pred. No. 1e-11; Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0  1 GCGSGGGGGGGGGGGGGGGGGGGGGGGGGGS 35           1 GCGSGGGGGGGGGGGGGGGGGGGGGGGGGGS 35
ID	AAWI8564 standard; protein; 312 AA. AAWI8564; AAWI8564; 27-MAR-1998 (first entry) Novel fusion protein KDTF7.



KW Kunitz type active site inhibitor domain; fusion protein; Factor VIIa;  
 KW FVIIa; linker; tissue factor; TF; treatment; procoagulant activity;  
 KM thrombosis; human.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX Key Location/Qualifiers  
 FT Domain 1..58  
 FT Domain /note="Kunitz type active site inhibitor domain"  
 FT Domain 59..93  
 FT Domain /note="linker peptide"  
 FT Domain 94..312  
 FT Domain /note="tissue factor domain"  
 XX WO9720939-A1.  
 XX 12-JUN-1997.  
 XX 22-NOV-1996; 96WO-US018756.  
 XX 01-DEC-1995; 95US-00566459.  
 XX 04-DEC-1995; 95US-00566800.  
 XX (GETH ) GENENTECH INC.  
 XX Kelley RF, Lazarus RA, Lee GF;  
 XX WPI, 1997-319785/29.  
 XX Factor VIIa active site inhibitor domain and tissue factor fusion protein  
 PT - used for inhibiting human tissue factor-FVIIa procoagulant activity in  
 PT mammals, to treat e.g. thrombosis and restenosis.  
 XX Example 2; Page; 58pp; English.  
 XX This is a novel fusion protein comprising a factor VIIa (FVIIa) Kunitz  
 CC type active site inhibitor domain sequence (AAW18553), a linker domain  
 CC sequence, and a tissue factor (TF) domain sequence which is a human  
 CC tissue factor variant hTFPA (AAW18556). This novel fusion protein has a  
 CC Kunitz domain linked at its C-terminus via the peptide linker to the N-  
 CC terminus of hTFPA. The novel fusion protein can be used for inhibiting  
 CC human tissue factor-FVIIa (TF-FVIIa) procoagulant activity in a mammal.  
 CC It can be used to treat TF-FVIIa related diseases like chronic  
 CC thromboembolic diseases or disorders associated with fibrin formation  
 CC including vascular disorders such as deep venous thrombosis, arterial  
 CC thrombosis, stroke, tumour metastasis, thrombolysis, arteriosclerosis and  
 CC restenosis following angioplasty. The novel fusion protein can also be  
 CC used to treat acute and chronic disorders such as inflammation, septic  
 CC shock, adult respiratory distress syndrome, septicemia, hypotension,  
 CC disseminated intravascular coagulopathy and other diseases. Fusion  
 CC proteins in which a FVIIa active site inhibitor and a tissue factor are  
 CC covalently linked are more potent inhibitors of FVIIa than the individual  
 CC components alone or a non-covalent mixture of the two. Note: This  
 CC sequence is not explicitly shown in the specification; it has been  
 CC created by combining a Kunitz domain sequence TF71-C (AAW18553), a linker  
 CC sequence given in Page 43 followed by a hTFPA sequence (AAW18556)  
 XX SQ Sequence 312 AA;  
 XX Query Match 100.0%; Score 196; DB 2; Length 312;  
 XX Best Local Similarity 100.0%; Pred. No. 3,9e-11;  
 XX Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AC AAB00158;  
 XX 08-FEB-2001 (first entry)  
 XX sCD4-SCFV(17b) HIV single chain antibody fusion protein.  
 DE sCD4-SCFV(17b) HIV single chain antibody fusion protein.  
 KW Fusion protein; HIV; human immunodeficiency virus; antibody; Fv; AIDS;  
 KW acquired immune deficiency syndrome; neutralisation; infection;  
 KW gene therapy; CD4; gp120; glycoprotein; resistance; vaccination;  
 KW binding domain; single chain antibody; chimera; chimeric protein.  
 OS Human immunodeficiency virus.  
 OS Synthetic.  
 XX WO200055207-A1.  
 XX 21-SEP-2000.  
 XX 16-MAR-2000; 2000MO-US006946.  
 XX 16-MAR-1999; 99US-0124681P.  
 XX (USSH ) US NAT INST OF HEALTH.  
 XX Berger EA, Del Castillo CM;  
 XX WPI, 2000-638183/61.  
 XX N-PSDB; AAW54045.  
 XX Novel neutralizing bispecific fusion proteins effective in viral such as  
 PT HIV neutralization, comprises two different binding domains, inducing-  
 PT binding domain and induced-binding domain functionally linked by linker.  
 XX Claim 39; Page 46-47; 55pp; English.  
 XX sCD4-SCFV(17b) is a neutralising bispecific fusion protein capable of  
 CC binding to two sites of its target protein. The protein comprises a first  
 CC binding domain capable of binding to an inducing site on the target  
 CC protein, a second binding domain capable of forming neutralising complex  
 CC with an induced epitope of the target protein and a linker connecting the  
 CC binding domains. sCD4-SCFV(17b) comprises a soluble CD4 fragment  
 CC (containing domains D1 and D2) fused to a single chain Fv portion of  
 CC antibody 17b via a linker. sCD4-SCFV(17b), its variant, analogue or  
 CC mimetic is used for inactivating gp120 protein of HIV, and for  
 CC neutralising HIV. It is also used for blocking and preventing the binding  
 CC of the viral or recombinant gp120 protein to soluble CD4 or lymphocyte  
 CC CD4 and for inhibiting HIV replication. The chimeric proteins is  
 CC therefore useful for treating HIV infection and also AIDS. It is are  
 CC particularly useful in the prevention of infection during or immediately  
 CC after HIV exposure (e.g., mother/infant transmission, post-exposure  
 CC prophylaxis, and as a topical inhibitor) and for providing long term  
 CC resistance to HIV infections and AIDS. Gene therapy is used to secrete  
 CC the bispecific protein at mucosal surfaces, such as the vaginal, rectal  
 CC or oral mucosa. The fusion proteins is highly potent, broadly cross-  
 CC reactive with neutralising antibody with high in vivo activity and no Fc-  
 CC mediated undesirable targeting properties. When the fusion protein is  
 CC substantially derived from human proteins, it has minimal immunogenicity  
 CC and toxicity in humans which is of great value in prevention of infection  
 CC during or immediately after HIV exposure  
 XX SQ Sequence 507 AA;  
 XX Query Match 100.0%; Score 196; DB 3; Length 507;  
 XX Best Local Similarity 100.0%; Pred. No. 5,7e-11;  
 XX Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 9  
 AAB00158  
 ID AAB00158 standard; protein; 507 AA.  
 XX

RESULT 10  
 AAB00159  
 ID AAB00159

ID AAB00159 standard; peptide; 39 AA.  
 XX AAB00159;  
 XX  
 DT 08-FEB-2001 (first entry)  
 DE Seven repeat linker.  
 KM Fusion protein; HIV; human immunodeficiency virus; antibody; Fv; AIDS;  
 KM acquired immune deficiency syndrome; neutralisation; infection;  
 KM gene therapy; CD4; gp120; glycoprotein; resistance; vaccination;  
 KM binding domain; single chain antibody; chimera; chimeric protein.  
 OS Synthetic.  
 PN W0200055207-A1.  
 PD 21-SEP-2000.  
 PF 16-MAR-2000; 2000MO-US006946.  
 PR 16-MAR-1999; 99US-0124681P.  
 PA (USSH ) US NAT INST OF HEALTH.  
 PI Berger EA, Del Caestillo CM;  
 PI WPI; 2000-638183/61.  
 DR N-PSDB; AAA54046.  
 PT Novel neutralizing bispecific fusion proteins effective in viral such as  
 PT HIV neutralization, comprises two different binding domains, inducing-  
 PT binding domain and induced-binding domain functionally linked by linker.  
 XX  
 PS Example 1; Page 48; 55pp; English.  
 CC sCD4-SCFv(17b) is a neutralising bispecific fusion protein capable of  
 CC binding to two sites of its target protein. The protein comprises a first  
 CC binding domain capable of binding to an inducing site on the target  
 CC protein, a second binding domain capable of forming neutralising complex  
 CC with an induced epitope of the target protein and a linker connecting the  
 CC binding domains. sCD4-SCFv(17b) comprises a soluble CD4 fragment  
 CC (containing domains D1 and D2) fused to a single chain Fv portion of  
 CC antibody 17b via a linker. sCD4-SCFv(17b), its variant, analogue or  
 CC mimetic is used for inactivating gp120 protein of HIV, and for  
 CC neutralising HIV. It is also used for blocking and preventing the binding  
 CC of the viral or recombinant gp120 protein to soluble CD4 or lymphocyte  
 CC CD4 and for inhibiting HIV replication. The chimeric proteins is  
 CC particularly useful for treating HIV infection and also AIDS. It is are  
 CC after HIV exposure (e.g., mother/infant transmission, post-exposure  
 CC prophylaxis, and as a topical inhibitor) and for providing long term  
 CC resistance to HIV infections and AIDS. Gene therapy is used to secrete  
 CC the bispecific protein at mucosal surfaces, such as the vaginal, rectal  
 CC or oral mucosa. The fusion proteins is highly potent, broadly cross-  
 CC reactive with neutralising antibody with high in vivo activity and no Fc-  
 CC mediated undesirable targeting properties. When the fusion protein is  
 CC substantially derived from human proteins, it has minimal immunogenicity  
 CC and toxicity in humans which is of great value in prevention of infection  
 CC during or immediately after HIV exposure. To generate the fusion protein  
 CC the full length CD4 coding sequence was excised from the plasmid pCB-3  
 CC using restriction enzymes StuI and SpeI. Synthetic oligonucleotides  
 CC (AAA54049, AAA54046) were ligated and used to regenerate the cleaved StuI  
 CC site and the next two bases of the CD4 cDNA and to produce an SpeI  
 CC overhang at the 3' end. The sequence also reconstructs the second domain  
 CC of CD4 and encodes the 37 amino acid intermediate linker  
 CC (Gly 4Ser)<sub>6</sub>Gly 4Thr 2Ser followed directly by the universal translation  
 CC termination signal. A BamHI site is included near the end of the third  
 CC (Gly 4Ser)<sub>6</sub> repeat to enable linkage to the 17b SCFv with the exact linker  
 CC sequence. The resulting intermediate plasmid was designated pCD1 but the  
 CC sCD4 linker insert was then subcloned into a pSCS3 derivative lacking a  
 CC BamHI site, forming pCD2. The 17b SCFv sequence was amplified from its  
 CC starting plasmid vector pmt del 0 using two primers (AAA54047, AAA54048).

CC The amplified fragment was digested with BamHI and SpeI and force-cloned  
 CC into pCD2 to generate the sCD4-SCFv(17b) fusion protein in a plasmid  
 CC designated pCD3  
 XX  
 SQ Sequence 39 AA;  
 Query Match 98.5%; Score 193; DB 3; Length 39;  
 Best Local Similarity 97.1%; Pred. No. 1,4e-11;  
 Matches 34; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGGSGGGSGGGSGGGSGGGSGGGSGGGSGGGSG 35  
 DB 3 GGGSGGGSGGGSGGGSGGGSGGGSGGGSGGGGT 37  
 RESULT 11  
 ID AB009677 standard; peptide; 40 AA.  
 XX  
 AC AB009677;  
 DT 02-JUL-2003 (first entry)  
 DE INF-beta-fragment of crystallisation fusion protein related linker #6.  
 XX  
 KM Interferon-beta; INF-beta; fragment of crystallisation; Fc; IgG4;  
 KM cytostatic; virucide; hepatotropic; antiinflammatory; tumour;  
 KM viral infection; interferon; INF; immunoglobulin Fc; hepatitis B;  
 KM lymphoma; leukaemia; linker.  
 XX  
 OS Synthetic.  
 PN US2003026779-A1.  
 PD 06-FEB-2003.  
 PF 03-DEC-2001; 2001US-00005438.  
 PR 15-OCT-1999; 99US-00418734.  
 PA (YUL/) YU L.  
 PA (CHAN/) CHANG T W.  
 PI Yu L, Chang TW;  
 XX WPI; 2003-417234/39.  
 DR  
 XX Treating tumors or viral infections, involves administering a hybrid  
 PT molecule comprising an interferon molecule joined at one end to a first  
 PT end of an immunoglobulin Fc fragment, with or without a linker.  
 XX  
 PS Example 2; Page 3; 9pp; English.  
 CC The invention describes a method of treating tumors or viral infections,  
 CC involving administering a hybrid molecule having an interferon (IFN)  
 CC molecule (I) or its variants, joined at one end to a first end of an  
 CC immunoglobulin Fc fragment (II), without any linker between (I) and (II)  
 CC or with a first linker between (I) and (II). The method is useful for  
 CC treating tumors and viral infections including hepatitis B. The IFN-Fc  
 CC hybrids have a much more longer half-life in vivo than the native  
 CC IFNalpha. They are suitable for treating tumors such as lymphomas and  
 CC leukemias, as these products have a long retention time in the  
 CC vasculature and will not penetrate undesired sites. The hybrids have been  
 CC shown to ablate tumors in an animal model. At equivalent molar dosage of  
 CC native IFN and the hybrid, the hybrid provides a several hundred fold  
 CC efficacy at the same dosage and less frequent administration. This is the  
 CC amino acid sequence of a linker that can be used to create an interferon-  
 CC beta (INF-beta)-IgG4 fragment of crystallisation (Fc) fusion protein  
 XX  
 SQ Sequence 40 AA;  
 Query Match 89.8%; Score 176; DB 6; Length 40;  
 Best Local Similarity 88.6%; Pred. No. 5.1e-10;



OS	Synthetic.
XX	
PN	WO2003024480-A2.
PD	
XX	27-MAR-2003.
PF	
XX	16-SEP-2002; 2002MO-IB004252.
PR	
XX	14-SEP-2001; 2001US-0318967P.
PA	
XX	(CYTO-) CYTOS BIOTECHNOLOGY AG.
PI	
XX	Bachmann MF, Storni T, Lechner F;
DR	WPI; 2003-363095/34.
XX	
PT	A composition, useful for enhancing an immune response against an antigen
PT	or a virus-like particle, enhancing anti-viral protection in an animal,
PT	or immunising or treating tumors or infectious diseases, e.g. viral
PS	infections.
XX	
PS	Disclosure, Page 64; 243pp; English.

CC The present invention describes a composition (C) for enhancing an immune  
CC response against an antigen or a virus-like particle in an animal. (C)  
CC comprises a virus-like particle (VLP) bound to at least one antigen, or a  
CC VLP capable of being recognised by the immune system of the animal. Also  
CC described: (1) enhancing an immune response against an antigen or a VLP  
CC in an animal comprising introducing (C) into the animal; (2) vaccines  
CC comprising (C) together with a pharmaceutical diluent, carrier or  
CC excipient; (3) immunising or treating an animal comprising administering  
CC the vaccine to the animal, or priming or boosting a T cell response in  
CC the animal by administering the vaccine; and (4) enhancing anti-viral  
CC protection in an animal comprising introducing (C) into the animal. (C)  
CC has cytostatic, virucide, antibacterial, antiparasitic, fungicide,  
CC antiallergic, immunosuppressive, antididiactive, antiinflammatory,  
CC antithyroid, antidiabetic, neuroprotective, nootropic, osteopathic,  
CC antirheumatic and antiarthritic activities. (C) or the vaccines can be  
CC used for enhancing an immune response against an antigen or a VLP in an  
CC animal, enhancing anti-viral protection in an animal, or immunising or  
CC treating tumours and infectious diseases such as viral, bacterial,  
CC parasitic or fungal infections. The vaccine compositions are also useful  
CC for preventing or treating allergies, drug addiction, graft-versus-host  
CC disease, Crohn's disease, Grave's disease, diabetes, multiple sclerosis,  
CC Alzheimer's disease, osteoporosis, rheumatoid arthritis, or inflammatory  
CC autoimmune disease. ACC69838 to ACC69852 and ABR56401 to ABR56509  
CC represent sequences used in the exemplification of the present invention  
XX Sequence 41 AA;

Query Match	87.8%	Score 172	DB 6	Length 41
Best Local Similarity	90.9%	Pred. No. 1,2e-09		
Matches	30	Conservative	0	Mismatches 3
				Indels 0
				Gaps 0
QY	1	GGGSGGGGSGGGGSGGGGSGGGGSGGGG	33	
Db	3	GGGGGGGGSGGGGSGGGGSGGGGSGGGGGG	35	

RESULT	15
ABR44505	
ID	ABR44505 standard; peptide; 41 AA
XX	
AC	ABR44505;
XX	
DT	25-JUL-2003 (first entry)

XX	Peptide linker #4.
DE	
XX	
KW	Immunostimulatory; virus-like particle; bacteriophage; HBV; LCMV;
KW	hepatitis B virus; lymphocytic choriomeningitis virus; vaccine;
KW	immunostimulant; cytostatic; anti-allergic; virucide; antibacterial;
KW	immune response; immunisation; allergy; tumour; breast cancer;
KW	neuroblastoma; leukaemia; viral disease; influenza; hepatitis; measles;
KW	chicken pox; bacterial infection; tuberculosis; pneumonia; syphilis.
XX	
OS	Synthetic.
XX	
PN	WO2003024481-A2.
XX	
PD	27-MAR-2003.
XX	
PE	16-SEP-2002; 2002WO-IB004132.
XX	
PR	14-SEP-2001; 2001US-0318994P.
XX	22-APR-2002; 2002US-0374145P.
XX	
PA	(CYTO-) CYTOS BIOTECHNOLOGY AG.
PA	(MAUR/) MAURER P.
PA	(TISS/) TISSOT A.
PA	(SCHW/) SCHWARZ K.
PA	(MEIJ/) MEIJERINK E.
PA	(LIPO/) LIPOWSKY G.
PA	(PUMP/) PUMPERTS P.
PA	(CIEL/) CIELENS I.
PA	(RENH/) RENHOFFA R.
XX	
PI	Maurer P, Tissot A, Schwarz K, Meijerink E, Lipowsky G;
PI	Pumpert P, Cielens I, Renhofa R, Bachmann WF, Storni T;
XX	
DR	WPI; 2003-354564/33.
XX	
XX	
PT	New compositions comprising immunostimulatory substances packaged into
PT	virus-like particles, useful as a vaccine for enhancing an immune
PT	response in animals, e.g. for treating or preventing allergies, tumors or
PT	viral infections.
XX	
PS	Disclosure; Page 75; 322p; English.

The present invention describes a composition (C) for enhancing an immune response in an animal. (C) comprises a virus-like particle (VLP), and an immunostimulatory substance. The immunostimulatory substance is bound to the VLP. Also described: (1) enhancing an immune response in an animal by introducing (C) into the animal; (2) producing (C) for enhancing an immune response in an animal; (3) vaccines comprising (C) together with a pharmaceutical diluent, carrier or excipient; and (4) immunising or treating an animal by: (a) administering the vaccine to the animal; (b) priming a T cell response in the animal by administering the vaccine; or (c) boosting a T cell response in the animal by administering the vaccine. (C) has immunostimulant, cytostatic, antiallergic, vitucide and antibacterial activities. (1) can be used in vaccines for enhancing an immune response in an animal, particularly a mammal or human. Specifically, (C) is useful for enhancing a B cell response, a T cell response, or a cytotoxic T-lymphocyte (CTL) response. (C) or a vaccine comprising (C) can also be used for immunising or treating an animal, e.g. humans, sheep, horses, cattle, pigs, dogs, cats, rats, birds, reptiles or fish. (C) is particularly useful in prophylactic or therapeutic vaccines against allergies, tumours (e.g. breast cancers, neuroblastoma, or leukaemia), viral diseases (e.g. influenza, hepatitis, measles or chicken pox), or bacterial infections (e.g. tuberculosis, pneumonia or syphilis). ACC69790 to ACC69815 and ABR44502 to ABR44612 represent sequences used in the exemplification of the present invention

Query Match	87.8%	Score 172;	DB 6;	Length 41;
Best Local Similarity	90.9%	Pred. No. 1.2e-09;		
Matches 30;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:29:56 ; Search time 4.53467 Seconds  
(without alignments)  
511.863 Million cell updates/sec

Title: US-09-936-702-2

Perfect score: 196

Sequence: 1 GGGGSGGGSGGGSGGGSGGGSGGGSGGGSG 35

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	196	100.0	40	3	US-08-815-190A-17
2	196	100.0	40	4	US-09-333-213-1
3	196	100.0	60	4	US-09-832-297A-12
4	196	100.0	100	4	US-09-411-067C-4
5	173.5	88.5	456	4	US-09-495-880A-11
6	172.5	88.0	738	3	US-08-864-038A-3
7	168	85.7	201	3	US-09-052-995-1
8	168	85.7	201	3	US-09-053-003-40
9	168	85.7	201	4	US-09-054-281-22
10	168	85.7	201	4	US-09-478-948-6
11	168	85.7	201	4	US-09-818-094-40
12	153	78.1	643	4	US-09-538-092-844
13	152	77.6	268	4	US-08-835-099A-9
14	152	77.6	268	3	US-09-157-349-9
15	150	76.5	190	3	US-09-046-985-15
16	150	76.5	190	3	US-09-474-743-15
17	148.5	75.8	647	2	US-08-770-761A-8
18	148.5	75.8	705	2	US-08-770-761A-7
19	146	74.5	266	4	US-09-032-523-7
20	146	74.5	266	3	US-09-802-633-7
21	144	73.5	33	3	US-09-215-212-3
22	143	73.0	420	2	US-08-845-998-8
23	143	73.0	420	3	US-09-206-537-8
24	143	73.0	420	3	US-09-430-854-8
25	141	71.9	345	4	US-09-657-013-112
26	140.5	71.7	316	4	US-09-538-092-997
27	140.5	71.7	495	3	US-08-828-741B-4

28	140.5	71.7	495	3	US-09-160-567-4	Sequence 4, Appl
29	140.5	71.7	495	4	US-09-710-299-4	Sequence 4, Appl
30	140.5	71.7	495	4	US-09-509-031-4	Sequence 4, Appl
31	140	71.4	25	4	US-09-254-832B-22	Sequence 22, Appl
32	140	71.4	112	4	US-09-543-681A-8319	Sequence 8319, Ap
33	140	71.4	979	4	US-09-538-092-990	Sequence 990, App
34	139	70.9	111	4	US-09-543-681A-8322	Sequence 8322, Ap
35	138	70.4	36	2	US-08-470-775-11	Sequence 11, Appl
36	138	70.4	36	3	US-08-469-318-50	Sequence 50, Appl
37	138	70.4	36	3	US-08-468-609A-50	Sequence 50, Appl
38	138	70.4	36	3	US-08-192-325B-50	Sequence 50, Appl
39	138	70.4	36	3	US-08-559-267A-11	Sequence 11, Appl
40	138	70.4	36	3	US-08-469-124-11	Sequence 11, Appl
41	138	70.4	36	3	US-08-446-872A-50	Sequence 50, Appl
42	138	70.4	36	4	US-08-762-227A-50	Sequence 50, Appl
43	138	70.4	36	4	US-09-510-238A-9	Sequence 9, Appl
44	138	70.4	36	5	PCT-US95-01185-50	Sequence 50, Appl
45	138	70.4	49	3	US-08-469-318-194	Sequence 194, App

#### ALIGNMENTS

RESULT 1  
US-08-815-190A-17  
Sequence 17, Application US/08815190A  
Patent No. 6046310  
GENERAL INFORMATION:  
APPLICANT: Queen, Cary L.  
APPLICANT: Schneider, William P.  
APPLICANT: Vasquez, Maximiliano  
TITLE OF INVENTION: Uses  
TITLE OF INVENTION: Fae Ligand Fusion Proteins and Their  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/815,190A  
FILING DATE: 11-MAR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/614,584  
FILING DATE: 13-MAR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 011823-00671005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 40 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Peptide  
FEATURES:  
NAME/KEY: Peptide  
LOCATION: 6..40  
OTHER INFORMATION: /notes="amino acid residues 6-40 may be  
OTHER INFORMATION: present or absent"  
US-08-815-190A-17







RESULT 9  
 US-09-054-281-22  
 Sequence 22, Application US/09054281  
 Patent No. 6444421  
 GENERAL INFORMATION:  
 APPLICANT: Chung, Jay H.  
 TITLE OF INVENTION: Methods for Detecting Intermolecular  
 TITLE OF INVENTION: Interactions In Vivo and In Vitro  
 NUMBER OF SEQUENCES: 22  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/054,281  
 FILING DATE: 02-APR-1998  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/080,234  
 FILING DATE: 03-APR-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/826,622

```

RESULT 10
US-09-478-948-6
/ Sequence 6, Application US/09478948
/ Patent No. 6465258
/ GENERAL INFORMATION:
/ APPLICANT: Shan, Bel
/ APPLICANT: Okamoto, Arthur Y.
/ APPLICANT: Tularik Inc.
/ TITLE OF INVENTION: FXR Receptor-Mediated Modulation of Cholesterol
/ TITLE OF INVENTION: Metabolism
/ FILE REFERENCE: 018781-001310US
/ CURRENT APPLICATION NUMBER: US/09/478,948
/ CURRENT FILING DATE: 2000-01-06
/ PRIOR APPLICATION NUMBER: US 60/115,249
/ PRIOR FILING DATE: 1999-01-07
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 6
/ LENGTH: 201
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:Flexible linker
/ NAME/KEY: MOD RES
/ LOCATION: (1)..(97)
/ OTHER INFORMATION: Gly at positions 1-97 may be present or absent
/ NAME/KEY: MOD RES
/ LOCATION: (105)..(201)
/ OTHER INFORMATION: Gly at positions 105-201 may be present or absent
US-09-478-948-6

Query Match      85.7%; Score 168; DB 4; Length 201;
Best Local Similarity 82.4%; Pred. No. 6,4e-09;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps

```



TELEFAX: 617-523-6440  
TELEX:  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-835-099A-9

Query Match 77.6%; Score 152; DB 2; Length 268;  
Best Local Similarity 63.6%; Pred. No. 2.2e-07;  
Matches 28; Conservative 1; Mismatches 5; Indels 10; Gaps 1;

Qy 1 GGGGGGGGGGGG-----GSGGGGGGGGGGGGGGGG 34  
Db 12 GGGGGGGGGGIGGGIGNVGLISGAGGGGGGGGGGGGG 55

## RESULT 14

US-09-157-349-9  
Sequence 9, Application US/09157349  
Patent No. 6068990  
GENERAL INFORMATION:  
APPLICANT: SHINTANI, Yasuaki  
APPLICANT: NISHI, Kazuo  
APPLICANT: KAWAMOTO, Tomohiro  
TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION  
TITLE OF INVENTION: AND USE  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP  
STREET: 130 Water Street  
City: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/157,349  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/835,099  
FILING DATE:  
APPLICATION NUMBER: 97105508.2  
FILING DATE: 03-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Reenick, David S  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 47342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
TELEX:  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-157-349-9

Query Match 77.6%; Score 152; DB 3; Length 268;  
Best Local Similarity 63.6%; Pred. No. 2.2e-07;  
Matches 28; Conservative 1; Mismatches 5; Indels 10; Gaps 1;

Qy 1 GGGGGGGGGGGG-----GSGGGGGGGGGGGGGGGG 34  
Db 12 GGGGGGGGGGIGGGIGNVGLISGAGGGGGGGGGGGGG 55

RESULT 15  
US-09-046-985-15  
Sequence 15, Application US/09046985  
Patent No. 6121236  
GENERAL INFORMATION:  
APPLICANT: Ben-Sasson, Shmuel A.  
TITLE OF INVENTION: NOVEL MULTIVALENT LIGANDS WHICH MODULATE  
TITLE OF INVENTION: ANGIOGENESIS  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
City: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/046,985  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: CMCC-614  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781) 861-6240  
TELEFAX: (781) 861-9540  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 190 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-046-985-15

Query Match 76.5%; Score 150; DB 3; Length 190;  
Best Local Similarity 40.2%; Pred. No. 2.5e-07;  
Matches 35; Conservative 0; Mismatches 0; Indels 52; Gaps 2;

Qy 1 GGGGGGGGGGGG-----GSGGGGGGGGGGGGGG 30  
Db 27 GGGGGGGGGGGGSDFTYRWRLSHRPKDLYSIVRRADGGGGGGGGGGGGKDF 86

Qy 31 -----GGGS 35  
Db 87 AYRRLSHRPKDLYSIVRRADGGGS 113  
Search completed: November 21, 2004, 13:44:44  
Job time : 5.53467 sec

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:42:32 ; Search time 15.3285 Seconds  
(without alignments)  
808.591 Million cell updates/sec

Title: US-09-936-702-2

Perfect score: 196

Sequence: 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 35

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1570615 seqs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	196	100.0	35	14	US-10-013-173-30
2	196	100.0	35	14	US-10-150-752-30
3	196	100.0	35	14	US-10-244-821-30
4	196	100.0	40	13	US-10-081-400-1
5	196	100.0	40	15	US-10-257-864A-137
6	196	100.0	40	15	US-10-257-864A-138
7	196	100.0	40	16	US-10-608-710-6
8	196	100.0	40	16	US-10-768-873-1
9	196	100.0	50	10	US-09-948-039-32
10	196	100.0	60	9	US-09-832-297A-12
11	196	100.0	60	10	US-09-833-203-18
12	182	92.9	211	16	US-10-437-963-126553
13	179.5	91.6	250	17	US-10-425-115-288960

14	176	89.8	40	14	US-10-005-438-9	Sequence 9, Appl
15	173.5	88.5	133	17	US-10-425-115-250884	Sequence 250884,
16	173.5	88.5	456	15	US-10-634-862-118955	Sequence 11, Appl
17	173	88.3	184	17	US-10-425-115-288955	Sequence 288955,
18	172	87.8	394	15	US-10-433-108-26	Sequence 26, Appl
19	172	87.8	302	15	US-10-433-108-25	Sequence 25, Appl
20	171	87.2	160	17	US-10-739-930-7816	Sequence 7816, Ap
21	168	85.7	30	10	US-09-833-203-19	Sequence 19, Appl
22	168	85.7	200	9	US-09-798-584-18	Sequence 18, Appl
23	168	85.7	200	9	US-09-967-624-19	Sequence 19, Appl
24	168	85.7	200	9	US-09-998-667-18	Sequence 18, Appl
25	168	85.7	200	10	US-09-921-159-14	Sequence 34, Appl
26	168	85.7	200	10	US-09-990-940-21	Sequence 21, Appl
27	168	85.7	200	10	US-09-989-981A-13	Sequence 13, Appl
28	168	85.7	200	10	US-09-850-948-29	Sequence 29, Appl
29	168	85.7	200	14	US-10-160-354-4	Sequence 4, Appl
30	168	85.7	200	14	US-10-026-021-8	Sequence 8, Appl
31	168	85.7	200	14	US-10-161-165-3	Sequence 3, Appl
32	168	85.7	200	14	US-10-160-663-3	Sequence 3, Appl
33	168	85.7	200	14	US-10-071-838-15	Sequence 15, Appl
34	168	85.7	200	14	US-10-094-417-25	Sequence 25, Appl
35	168	85.7	200	14	US-10-188-405-13	Sequence 13, Appl
36	168	85.7	200	14	US-10-273-575-29	Sequence 29, Appl
37	168	85.7	200	14	US-10-233-098-5	Sequence 5, Appl
38	168	85.7	200	14	US-10-245-850-3	Sequence 3, Appl
39	168	85.7	200	14	US-10-293-582-27	Sequence 27, Appl
40	168	85.7	200	14	US-10-100-818-14	Sequence 14, Appl
41	168	85.7	200	14	US-10-237-467-18	Sequence 18, Appl
42	168	85.7	200	14	US-10-321-204-54	Sequence 54, Appl
43	168	85.7	200	14	US-10-179-766-12	Sequence 12, Appl
44	168	85.7	200	14	US-10-123-568-4	Sequence 4, Appl
45	168	85.7	200	14	US-10-123-731-10	Sequence 10, Appl

## ALIGNMENTS

RESULT 1  
US-10-013-173-30  
; Sequence 30, Application US/10013173  
; Publication No. US20030095977A1  
; GENERAL INFORMATION:  
; APPLICANT: Goshorn, Stephen C.  
; APPLICANT: Graves, Scott Stoll  
; APPLICANT: Schultz, Joanne Elaine  
; APPLICANT: Lin, Yukang  
; APPLICANT: Sanderson, James A.  
; APPLICANT: Reno, John M.  
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 690022.547C1  
; CURRENT APPLICATION NUMBER: US/10/013.173  
; CURRENT FILING DATE: 2001-12-07  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 35  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Linker sequence  
US-10-013-173-30

Query Match 100.0%; Score 196; DB 14; Length 35;  
Best Local Similarity 100.0%; Pred. No. 2.6e-09;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 35  
| | | | | | | | | | | | | | | | | | | | |  
DB 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 35  
| | | | | | | | | | | | | | | | | | | | |

RESULT 2

```

US-10-150-762-30
; Sequence 30, Application US/10150762
; Publication No. US20030103948A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott S.
; APPLICANT: Schultz, Joanne E.
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; APPLICANT: Deaslyne, Erica A.
; TITLE OF INVENTION: SREBP-DEPENDENT EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C2
; CURRENT APPLICATION NUMBER: US/10/150,762
; CURRENT FILING DATE: 2002-05-17
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker sequence
US-10-150-762-30

```

Query Match	100.0%	Score 196	DB 14	Length 35
Best Local Similarity	100.0%	Pred. No. 2.6e-09		
Matches 35, Conservative	0	Mismatches 0	Indels 0	Gaps 0

[illegible]

```

RESULT 3
US-10-244-821-30
?
? Sequence 30, Application US/10244821
? Publication No. US20030143233A1
? GENERAL INFORMATION:
? APPLICANT: Goshorn, Stephen Charles
? APPLICANT: Graves, Scott Scoll
? APPLICANT: Schultz, Joanne Elaine
? APPLICANT: Liu, Yikang
? APPLICANT: Sanderson, James Allen
? APPLICANT: Reno, John M.
? APPLICANT: Deareylene, Erica A.
? TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
? FILE REFERENCE: 696002.547C3
? CURRENT APPLICATION NUMBER: US/10/244,821
? CURRENT FILING DATE: 2002-09-16
? NUMBER OF SEQ ID NOS: 92
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 30
? LENGTH: 35
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Linker sequence
US-10-244-821-30

```

Query Match	100.0%	Score 196	DB 14	Length 35
Best Local Similarity	100.0%	Pred. No. 2.6e-09		
Matches 35, Conservative	0	Mismatches 0	Indels 0	Gaps 0

Qy 1 35  
Db 1 35

RESULT 4  
US-10-081-400-1

```

; Sequence 1, Application US/10081400
; Publication No. US2002015598A1
; GENERAL INFORMATION:
; APPLICANT: Young, Michael
; APPLICANT: Meade, Harry
; APPLICANT: Krane, Ian
; TITLE OF INVENTION: ERYTHROPOIETIN ANALOGY-HUMAN SERUM ALBUMIN FUSION
; FILE REFERENCE: 10275/041001
; CURRENT APPLICATION NUMBER: US/10/081,400
; CURRENT FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 09/333,213
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated linker sequence, subsets 2 through 8 (each
; OTHER INFORMATION: consisting of a repetition of the first five amino acids) encompass
; OTHER INFORMATION: positions 6 through 40 may be absent or present
US-10-081-400-1

```

Query Match	100.0%	Score 196;	DB 13;	Length 40;
Best Local Similarity	100.0%	Pred. No. 2.9e-09;		
Matches 35; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

```

? RESULT 5
? US-10-257-864A-137
? Sequence 137, Application US/10257864A
? Publication No. US20040058393A1
? GENERAL INFORMATION:
? APPLICANT: FUKUSHIMA, NAOSHI
? APPLICANT: TSUCHIYA, MASAYUKI
? APPLICANT: OH-EDA, MASAYOSHI
? APPLICANT: UNO, SHINSUKE
? APPLICANT: KIKUCHI, YASUFUMI
? APPLICANT: GOTEMBA-SHI, OHTOMO
? TITLE OF INVENTION: AGONIST ANTIODIES
? FILE REFERENCE: 065678-0107
? CURRENT APPLICATION NUMBER: US/10/257,864A
? CURRENT FILING DATE: 2003-07-24
? PRIOR APPLICATION NUMBER: PCT/JP01/03288
? PRIOR FILING DATE: 2001-04-17
? PRIOR APPLICATION NUMBER: PCT/JP01/01912
? PRIOR FILING DATE: 2001-03-12
? PRIOR APPLICATION NUMBER: JP2000-115246
? PRIOR FILING DATE: 2000-04-17
? PRIOR APPLICATION NUMBER: JP2000-321821
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: JP2000-321822
? PRIOR FILING DATE: 2000-10-20
? NUMBER OF SEQ ID NOS: 138
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 137
? LENGTH: 40
? TYPE: PPT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Synthetic
? OTHER INFORMATION: peptide
? FEATURE:
? NAME/KEY: PEPTIDE
? LOCATION: (1)..(40)
? OTHER INFORMATION: This peptide may range from 1-40 amino acids;
? OTHER INFORMATION: refer to the specification as filed for preferred
? OTHER INFORMATION: embodiments

```

```

; SEQ ID NO 32
:
: LENGTH: 50

```

```

/
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/
/ TITLE OF INVENTION: Plants
/
/ FILE REFERENCE: 38-21(53222)B
/
/ CURRENT APPLICATION NUMBER: US/10/425,115
/
/ CURRENT FILING DATE: 2003-04-28
/
/ NUMBER OF SEQ ID NOS: 369326
/
/ SEQ ID NO 288960
/
/ LENGTH: 250
/
/ TYPE: prt
/
/

```





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:29:16 ; Search time 3.83212 Seconds  
(without alignments)  
878.780 Million cell updates/sec

Title: US-09-936-702-2  
Perfect score: 196  
Sequence: 1 GGGSGGGSGGGSGGGSGGGSGGGSGGGSGGGSGGS 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	162.5	82.9	207	2 T07381	glycine-rich prote
2	162	82.7	221	2 T04592	glycine-rich cell
3	159.5	81.4	165	1 KNRZG1	glycine-rich cell
4	159	81.1	136	2 T29282	hypothetical prote
5	155	79.1	592	2 E82759	endo-1,4-beta-gluc
6	154.5	78.8	1226	2 T24045	hypothetical prote
7	154	78.6	266	1 CTRBL	calpain (EC 3.4.22
8	153	78.1	643	1 KHRU2	keratin 1, type II
9	152	77.6	268	1 CTRBL	calpain (EC 3.4.22
10	150.5	76.8	255	2 B84777	hypothetical prote
11	150	76.5	106	2 F84797	hypothetical prote
12	149.5	76.3	440	2 S71795	transcription fact
13	149	76.0	302	2 C84700	hypothetical prote
14	148.5	75.8	396	2 T49109	glycine-rich prote
15	148	75.5	259	2 T15126	hypothetical prote
16	147	75.0	481	2 A35628	loricrin - mouse
17	146.5	74.7	393	2 T20268	hypothetical prote
18	146	74.5	266	1 CTRGL	calpain (EC 3.4.22
19	145.5	74.2	271	2 S34666	glycine-rich prote
20	145	74.0	203	1 JQ1061	glycine-rich prote
21	145	74.0	385	2 T20410	hypothetical prote
22	145	74.0	404	2 S54729	RNA-binding protei
23	145	74.0	569	1 KRMSE1	keratin, 59k type
24	144.5	73.7	291	1 S31415	glycine-rich prote
25	144.5	73.7	581	1 KRMSE2	keratin, type II c
26	143.5	73.2	384	1 A26099	glycine-rich cell
27	143	73.0	139	2 S31443	glycine-rich RNA-b
28	143	73.0	158	2 T05254	probable RNA-bindi
29	143	73.0	171	2 H84709	probable glycine-r

30	143	73.0	420	2 I59234	octamer binding tr
31	142	72.4	408	2 S57483	glycin-rich protei
32	142	72.4	420	2 A49642	transcription fact
33	142	72.4	561	2 A31994	keratin 10, type I
34	141.5	72.2	681	2 AB2155	hypothetical prote
35	141	71.9	183	2 PN0109	keratin-like prote
36	141	71.9	839	2 B96576	hypothetical prote
37	140.5	71.7	316	1 A38743	loricrin - human
38	140	71.4	570	2 S07330	keratin, epidermal
39	140	71.4	979	2 A35913	regulatory factor
40	139	70.9	252	1 S01821	glycine-rich prote
41	139	70.9	284	2 S74256	homeotic protein s
42	139	70.9	330	2 S74255	homeotic protein s
43	138.5	70.7	212	2 A44994	eggshell protein 1
44	138.5	70.7	561	2 T22917	probable ATP-depen
45	138	70.4	424	1 Z3BPF1	coat protein A pre

## ALIGNMENTS

```

RESULT 1
T07381
glycine-rich protein Tfms - tomato
C:Species: Lycopersicon esculentum (tomato)
C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C:Accession: T07381
R:Santino, C.G.; Stanford, G.L.; Conner, T.W.
Plant Mol. Biol. 33, 405-416, 1997
A>Title: Developmental and transgenic analysis of two tomato fruit enhanced genes.
A:Reference number: Z16000; MUID:97201476; PMID:9049262
A:Accession: T07381
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-207 <SAS>
A:Cross-references: UNIPROT:Q43522; EMBL:X95262; NID:g1166449; PIDN:CAA64559.1; PID:g116
C:Genetics:
A:Gene: Tfms
C:Superfamily: hydroxyproline-rich glycoprotein

Query Match      82.9% Score 162.5; DB 2; Length 207;
Best Local Similarity 81.1% Pred. No. 4.6e-07;
Matches 30; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY      1 GGGSGGGGS---GGGSGGGSGGGSGGGSGGGSGGGG 34
DB      84 GGSAGGGSGTGGGGSGGGGGGGGGGGGGGGGGG 120

RESULT 2
T04592
glycine-rich cell wall structural protein homolog F23E13.120 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T04592
R:Bevan, M.; Halbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Jesse, T.
submitted to the Protein Sequence Database, March 1998
A:Reference number: Z15378
A:Accession: T04592
A:Molecule type: DNA
A:Residues: 1-221 <BEV>
A:Cross-references: UNIPROT:O65514; EMBL:AL022141
A:Experimental source: cultivar Columbia; BAC clone F23E13
C:Genetics:
A:Map position: 4
A>Note: F23E13.120

Query Match      82.7% Score 162; DB 2; Length 221;
Best Local Similarity 69.8% Pred. No. 5.3e-07;
Matches 30; Conservative 0; Mismatches 5; Indels 8; Gaps 1;

QY      1 GGGSG-----GGGSGGGSGGGSGGGSGGGSGGGSGGS 35

```



calpain (EC 3.4.22.17) small chain - rabbit  
 N:Alternate names: calcium-activated neutral proteinase (CANP); calpain light chain; cal  
 C:Species: *Oryctolagus cuniculus* (domestic rabbit)  
 C>Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 09-Jul-2004  
 C:Accession: A24816  
 J:Emori, Y.; Kawasaki, H.; Imajoh, S.; Kawashima, S.; Suzuki, K.  
 J. Biol. Chem. 261, 9472-9476, 1986  
 A>Title: Isolation and sequence analysis of cDNA clones for the small subunit of rabbit  
 A:Reference number: A24816; MUID:86250903; PMID:3013892  
 A:Accession: A24816  
 A:Molecule type: mRNA  
 A:Residues: 1-266 <EM>  
 A:Cross-references: UNIPROT:P06813; GB:M13364; NID:G164875; PIDN:AAA81565.1; PID:G164876  
 C:Complex: heterodimer of L (large) and S (small) chains  
 C:Function:  
 A:Description: catalyzes the hydrolysis of peptides  
 A>Note: cleaves preferentially after tyrosine, methionine, or arginine residues and befo  
 C:Superfamily: calpain small chain; calmodulin repeat homology  
 C:Keywords: calcium binding; cysteine proteinase; duplication; EF hand; heterodimer; hyd  
 F:1-54/Domain: glycine-rich <GLY>  
 F:94-125/Domain: calmodulin repeat homology <EF1>  
 F:137-169/Domain: calmodulin repeat homology <EF2>  
 F:170-199/Domain: calmodulin repeat homology <EF3>  
 F:202-234/Domain: calmodulin repeat homology <EF4>  
 F:235-266/Domain: calmodulin repeat homology <EF5>

Query Match 78.6%; Score 154; DB 1; Length 266;  
 Best Local Similarity 62.2%; Pred. No. 2.6e-06;  
 Matches 28; Conservative 3; Mismatches 4; Indels 10; Gaps 1;

Qy 1 GGGGSGGGSGGGG-----GGSGGGSGGGSGGGSGGGG 35  
 Db 11 GGGGSGGGGGLGGGLGNVLGGLISGAGGGGGGGGGGAGGAGGT 55

## RESULT 8

KRMU2

keratin 1, type II, cytoskeletal - human  
 N:Alternate names: 67K type II epidermal keratin; cytokeratin 1  
 C:Species: *Homo sapiens* (man)  
 C>Date: 04-Dec-1986 #sequence\_revision 22-Oct-1999 #text\_change 10-Dec-1999  
 C:Accession: A22940; A02950; A43342  
 R:Johnson, L.D.; Idler, W.W.; Zhou, X.M.; Roop, D.R.; Steinert, P.M.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 1896-1900, 1985  
 A:Reference number: A22940; MUID:85166239; PMID:2580302  
 A:Accession: A22940  
 A:Molecule type: DNA  
 A:Residues: 1-643 <J0H>  
 A:Cross-references: GB:M1215; GB:M11845; GB:M11846; NID:G1843461  
 A>Note: translation of initiator Met 18 not shown  
 R:Steinert, P.M.; Parry, D.A.D.; Idler, W.W.; Johnson, L.D.; Steven, A.C.; Roop, D.R.  
 U. Biol. Chem. 260, 7142-7149, 1985  
 A>Title: Amino acid sequences of mouse and human epidermal type II keratins of M-r 67,00  
 late filament subunits.  
 A:Reference number: A92535; MUID:85207740; PMID:2581964  
 A:Accession: A02950  
 A:Molecule type: mRNA  
 A:Residues: 151-183, 'K', 185-199, 'M', 201-204, 'K', 206-236, 'S', 238-239, 'R', 241-356, 'Y', 358-  
 'S', 638-643 <STE>  
 A:Cross-references: GB:M10938; NID:G186787; PIDN:AAA36153.1; PID:G386854  
 A:Experimental source: tissue neonatal foreskin  
 A>Note: the authors translated the codon CUG for residue 476 as Met  
 R:Chipev, C.C.; Kotze, B.P.; Markova, N.; Bale, S.J.; Digiovanna, J.J.; Compton, J.G.; S  
 Cell 70, 821-828, 1992  
 A>Title: A leucine----proline mutation in the H1 subdomain of keratin 1 causes epidermol  
 A:Reference number: A43342; MUID:92286601; PMID:1381288  
 A:Accession: A43342  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 144-146, 'P', 148-159, 'P', 161-183, 'K', 185-186 <CHI>  
 A:Cross-references: GB:M98776; GB:M1215; GB:M11845; GB:M11846; NID:G1843461  
 A>Note: sequence extracted from NCBI backbone (NCBI:112784)  
 C:Comment: The cytoskeletal and microfibrillar keratins are classified into two types, b

atin IF protein subunit appears to be a heterotrimer of two type I and two type II pro

C:Comment: Keratin 1 is expressed in terminally differentiating epidermis.  
 C:Genetics:  
 A:Gene: GDB:KRT1  
 A:Cross-references: GDB:128198; OMIM:139350  
 A:Map position: 12q11-12q13  
 A>Note: defects in this gene may result in epidermolytic hyperkeratosis  
 C:Complex: heterotrimer of two type I, usually keratin 10 (see PIR:KRU0), and two typ  
 C:Superfamily: cytoskeletal keratin  
 C:Keywords: coiled coil; heterotrimer; intermediate filament  
 F:4-179/Domain: head <HED>  
 F:143/Region: E1 and V1 subdomains  
 F:14-179/Region: H1 subdomain  
 F:180-492/Domain: rod <ROD>  
 F:180-214/Region: coil 1A  
 F:215-226/Region: linker 1A  
 F:227-327/Region: coil 1B  
 F:328-344/Region: linker 12  
 F:345-363/Region: coil 2A  
 F:364-371/Region: linker 2  
 F:372-492/Region: coil 2B  
 F:430/Region: stutler  
 F:493-643/Domain: tail <END>  
 F:493-512/Region: H2 subdomain  
 F:513-643/Region: V2 and E2 subdomains

Query Match 78.1%; Score 153; DB 1; Length 643;  
 Best Local Similarity 82.4%; Pred. No. 6.3e-06;  
 Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GGGGSGGGSGGGSGGGSGGGSGGGSGGGG 34  
 Db 93 GGGGFGGGGFGGGGFGGGGFGGGGFGGGGFGGGG 126

## RESULT 9

CITUU

calpain (EC 3.4.22.17) small chain - human  
 N:Alternate names: calcium-activated neutral proteinase (CANP)  
 C:Species: *Homo sapiens* (man)  
 C>Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 09-Jul-2004  
 C:Accession: A26107; A23650  
 R:Miyaake, S.; Emori, Y.; Suzuki, K.  
 Nucleic Acids Res. 14, 8805-8817, 1986  
 A>Title: Gene organization of the small subunit of human calcium-activated neutral prote  
 A:Reference number: A93648; MUID:87066759; PMID:3024120  
 A:Accession: A26107  
 A:Molecule type: DNA  
 A:Residues: 1-268 <MTY>  
 A:Cross-references: UNIPROT:P04632; GB:M31502  
 R:Ohno, S.; Emori, Y.; Suzuki, K.  
 Nucleic Acids Res. 14, 5559, 1986  
 A>Title: Nucleotide sequence of a cDNA coding for the small subunit of human calcium-dep  
 A:Reference number: A93631; MUID:86286563; PMID:3016651  
 A:Accession: A23650  
 A:Molecule type: mRNA  
 A:Residues: 1-268 <OHX>  
 A:Cross-references: EMBL:X04106; NID:G35327; PIDN:CAA27726.1; PID:G35328  
 C:Genetics:  
 A:Gene: GDB:CANP4  
 A:Cross-references: GDB:119752; OMIM:114170  
 A:Map position: 19pter-19pter  
 A:Introns: 70/2; 81/3; 111/3; 131/1; 152/3; 175/3; 202/1; 241/1; 260/3  
 C:Complex: heterodimer of L (large) and S (small) chains  
 C:Function:  
 A:Description: catalyzes the hydrolysis of peptides  
 A>Note: cleaves preferentially after tyrosine, methionine, or arginine residues and befo  
 C:Superfamily: calpain small chain; calmodulin repeat homology  
 C:Keywords: calcium binding; cysteine proteinase; duplication; EF hand; heterodimer; hyd  
 F:1-56/Domain: glycine-rich <GLY>  
 F:96-127/Domain: calmodulin repeat homology <EF1>  
 F:133-171/Domain: calmodulin repeat homology <EF2>  
 F:172-201/Domain: calmodulin repeat homology <EF3>



A:Experimental source: cultivar Columbia; BAC clone FIN20

C:Genetics:

A:Gene: ATSP:AT4g22020

A:Map position: 4

C:Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match

Best Local Similarity 75.8%; Score 148.5; DB 2; Length 396;

Matches 29; Conservative 0; Mismatches 6; Indels 19; Gaps 1;

QY 1 GGGSGGGSGG-----GGSGGGSGGGSGGGSGGGSG 35

DB 251 GGGGGGGGGGKVGCGYGHSGFGCGVGFNSGGGGGGGGGGGGGGG 304

#### RESULT 15

T15126

hypothetical protein T20B6.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C:Accession: T15126

R:Beck, C.; Wamsley, P.

Submitted to the EMBL Data Library, April 1997

A:Description: The sequence of C. elegans cosmid T20B6.

A:Reference number: Z18297

A:Accession: T15126

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-259 <BEC>

A:Cross-references: UNIPROT:002049; EMBL:AF000193; NID:G1946986; PID:G1946988; PIDN:AAB5

A:Experimental source: strain Bristol N2; clone T20B6

C:Genetics:

A:Gene: CESP:T20B6.3

A:Map position: 3

A:Introns: 9/2; 231/1

C:Superfamily: Phaseolus glycine-rich protein 1.0

Query Match

Best Local Similarity 75.5%; Score 148; DB 2; Length 259;

Matches 28; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 1 GGGSGGGSGG--GGSGGGSGGGSGGGSGGGSGGG 34

DB 184 GGGGFGGGMGYGGMGGGGYGGGGGGGGYGGGG 219

Search completed: November 21, 2004, 13:43:27  
Job time : 4.83212 secs

**THIS PAGE BLANK (uspto)**



1	172.5	88.0	738	2	0023402
2	169	86.2	551	2	Q75SHG8
3	169	86.2	551	2	AAR89849
4	162.5	82.9	207	2	Q43522
5	162	82.7	221	2	Q65514
6	161	82.1	192	2	Q92P87
7	160	81.6	233	2	Q92NUT
8	159.5	81.4	165	1	GKPI_ORYSA
9	159	81.1	136	2	Q18444
10	158	80.6	236	2	Q8RU50
11	157	80.1	698	2	Q8OKX8
12	156	79.6	397	2	Q6S9V4
13	156	79.6	397	2	AAR21812
14	156	79.6	527	2	Q6S9V3
15	156	79.6	527	2	AAR23813
16	155	79.1	592	2	Q9PF60
17	155	79.1	1615	2	Q7RRQM4
18	154.5	78.8	1160	2	Q7YXX8
19	154	78.6	161	2	Q95UX3
20	154	78.6	162	2	Q95UX5
21	154	78.6	163	2	Q95UN6
22	154	78.6	163	2	Q95UX4
23	154	78.6	164	2	Q95NP2
24	154	78.6	165	2	Q9GP44
25	154	78.6	165	2	Q95NR6
26	154	78.6	165	2	Q95UX2
27	154	78.6	266	1	CSSL_RABIT
28	154	78.6	284	2	Q73EF9
29	154	78.6	284	2	AAS39446
30	154	78.6	340	2	Q81E8
31	154	78.6	697	2	OGRINW

32	153	78.1	643	1	K2C1	HUMAN
33	152	78.1	644	2	AH63	997
34	152	77.6	167	2	O9LQZ8	
35	152	77.6	268	1	CSS1	HUMAN
36	152	77.6	2359	2	O8I519	
37	151	77.0	337	2	O9ET75	
38	151	77.0	718	2	O9ITR1	
39	150.5	76.8	255	2	O9S1H2	
40	150.5	76.8	774	2	O8E88A	
41	150.5	76.8	774	2	CAD88219	
42	150.5	76.8	774	2	CAD88221	
43	150	76.5	106	2	F7XJXP7	
44	149.5	76.3	440	1	PXGA	CHICK
45	149	76.0	267	2	Q6D1T2	
					Q6d1t2	xenopus tro

## ALIGNMENTS

```

RESULT 1
002402 ID 002402 PRELIMINARY; PRT; 738 AA.
AC 002402;
DT 01-JUL-1997 (TEMBirel. 04, Created)
DT 01-JUL-1997 (TEMBirel. 04, last sequence update)
DT 01-OCT-2003 (TEMBirel. 25, last annotation update)
DE Insoluble protein.
OS
NCBI_taxid=50426;
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pterioida;
OC Pterioidea; Pteridiidae; Pinctada.
CX
NCBI_Taxid=50426;
RN
RM
RX MEDLINE=97320490; PubMed=9177341;
RA Sudo S., Fujikawa T., Nakakura T., Ohkubo T., Sakaguchi K., Tanaka M.
RA Nakashima K., Takahashi T.;
RT "Structures of mollusc shell framework proteins.";
RL Nature 387:563-564(1997) .
DB EMBL; D86074; BA020466.1; -.
SQ SEQUENCE 738 AA; 61723 NM; FDF984139BF3BA59 CRC64;

```

Query Match: 88.0%; Score 172.5; DB 2; Length 738;

Matches 33; Conservative 0; Mismatches 2; Indels 7; Gaps 1

[illegible][illegible]

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

Q75HG8

AC Q75HG8;

DT 05-JUL-2004 (Tremblérel, 27, Last sequence update)

DE Putative prohibitin.

*Oryza sativa* (japonica cultivar-group).

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae

NCBI\_TaxID=39947;

RP SEQUENCE FROM N.A.

RA Overton II L.L., Tsicrin T., Kim M.M., Bera J.J., Jin S.S.,

RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblum T.V.,  
"O'Connell T.T." "Utterback T.T."

RA white O., Salzberg S.L., Fraser C.M.;  
 calculated (from 2003) to the turn/century/cent



[illegible]

CC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
CX NCBI\_TaxId=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21366234; PubMed=11474104;  
RC STRAIN=1021;  
RA Gilbert F., Finan T.M., Long S.R., Puhler A., Abola P., Ampe F.,  
BA Bailly-Hublier F., Barnett M.U., Becker A., Bolstad P., Bothe G.,  
RA Bouty M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,  
RA Cowie A., Davis R.W., Dreano S., Federpiel N.A., Fisher R.F.,  
RA Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M.,  
RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,  
RA Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Laetare V.,  
MA Mauby D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,  
RA Rameberger U., Surzycki K., Thebaud P., Vandebol M.,  
RA Vorholster F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;  
RL "The composite genome of the legume symbiont Sinorhizobium meliloti";  
RN science 293:668-672(2001).  
[2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=1021;  
RC MEDLINE=21396507; PubMed=11481430;  
RA Capela D., Bailly-Hublier F., Gouzy J., Bothe G., Ampe F., Batut J.,  
BA Bolstad P., Becker A., Bouty M., Cadieu E., Dreano S., Gloux S.,  
RA Godrie T., Goffeau A., Kahn D., Kiss E., Letare V., Mauby D.,  
RA Renard T., Portetelle D., Puhler A., Purnelle B., Rameberger U.,  
RT Pohl K., Thebaud P., Vandebol M., Weidner S., Gilbert F.;  
RL "Analysis of the chromosome sequence of the legume symbiont  
Sinorhizobium meliloti strain 1021";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
KW EMBL; AL591789; CAC46652.1;-  
SQ COMPLETE PROTEOME.  
SQ SEQUENCE 233 AA; 22513 MW; ED22171714B58E55 CRC64;

Query Match 81.6%; Score 160; DB 2; Length 233;  
Best Local Similarity 79.4%; Pred. No. 4,8e-06;  
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

DY 1 GGCGSGGCGSGCGGSGCGGSGCGGSGCGGSGCGG 34  
DB 103 GGCGSKGGGCGVCGGCGGCGGCGGCGGCGGCGG 136

RESULT 8  
GRP1\_ORYSA STANDARD; PRT; 165 AA.

ID GRP1\_ORYSA  
AC P25074;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Glycine-rich cell wall structural protein 1 precursor.  
GN Name=GRP-1;  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatopsida; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzoideae; Oryza.  
OX NCBI\_TaxId=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Indica / IR36;  
RX MEDLINE=91370862; PubMed=1716496;  
RA Lei M., Wu R.;  
RT "A novel glycine-rich cell wall protein gene in rice.";  
RL Plant Mol. Biol. 16:187-198(1991).  
CC -!- FUNCTION: Responsible for plasticity of the cell wall (Potential).  
CC -!- SUBCELLULAR LOCATION: Cell wall (Potential).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/



Db 471 GGGGGGTGGAGGGGGGGGGGGGGGGGGGGG 505

## RESULT 12

06S9V4 PRELIMINARY; PRT; 397 AA.  
AC 06S9V4;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE DSXF.  
GN Name=dax;  
OS Musca domestica (House fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Muscoidea;  
OC Muscidae; Musca.  
OX NCBI\_TaxID=7370;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=14673649;  
RA Hediger M., Burghardt G., Siegenthaler C., Buser N.,  
RA Hilfinger-Kleiner D., Dubendorfer A., Bopp D.;  
RT "Sex determination in Drosophila melanogaster and Musca domestica  
converges at the level of the terminal regulator doublesex.";  
RL Dev. Genes Evol. 214:29-42(2004).  
DR EMBL; AY461853; AAR23812.1; -  
DR InterPro; IPR001275; DM\_DNA-binding.  
DR Pfam; PF00751; DM; 1.  
DR SMART; SM00301; DM; 1.  
DR PROSITE; PS40000; DM\_DOMAIN\_1; 1.  
DR PROSITE; PS50809; DM\_DOMAIN\_2; 1.  
SQ SEQUENCE 397 AA; 41701 MW; 38B49CED6931F8C4 CRC64;

Query Match 79.6%; Score 156; DB 2; Length 397;  
Best Local Similarity 76.5%; Pred. No. 1.6e-05;  
Matches 26; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 34  
Db 192 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 225

## RESULT 13

AAR23812 PRELIMINARY; PRT; 397 AA.  
AC AAR23812;  
DT 02-MAR-2004 (TREMBlrel. 27, Created)  
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
DE DSXF.  
GN DSX.  
OS Musca domestica (House fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Muscoidea;  
OC Muscidae; Musca.  
OX NCBI\_TaxID=7370;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=14673649;  
RA Hediger M., Burghardt G., Siegenthaler C., Buser N.,  
RA Hilfinger-Kleiner D., Dubendorfer A., Bopp D.;  
RT "Sex determination in Drosophila melanogaster and Musca domestica  
converges at the level of the terminal regulator doublesex.";  
RL Dev. Genes Evol. 214:29-42(2004).  
DR EMBL; AY461853; AAR23812.1; -  
SQ SEQUENCE 397 AA; 41701 MW; 38B49CED6931F8C4 CRC64;

Query Match 79.6%; Score 156; DB 2; Length 397;  
Best Local Similarity 76.5%; Pred. No. 1.6e-05;  
Matches 26; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 34

Db 192 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 225

## RESULT 14

06S9V3 PRELIMINARY; PRT; 527 AA.  
AC 06S9V3;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE DSXM.  
GN Name=dax;  
OS Musca domestica (House fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Muscoidea;  
OC Muscidae; Musca.  
OX NCBI\_TaxID=7370;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=14673649;  
RA Hediger M., Burghardt G., Siegenthaler C., Buser N.,  
RA Hilfinger-Kleiner D., Dubendorfer A., Bopp D.;  
RT "Sex determination in Drosophila melanogaster and Musca domestica  
converges at the level of the terminal regulator doublesex.";  
RL Dev. Genes Evol. 214:29-42(2004).  
DR EMBL; AY461854; AAR23813.1; -  
DR InterPro; IPR001275; DM\_DNA-binding.  
DR Pfam; PF00751; DM; 1.  
DR SMART; SM00301; DM; 1.  
DR PROSITE; PS40000; DM\_DOMAIN\_1; 1.  
DR PROSITE; PS50809; DM\_DOMAIN\_2; 1.  
SQ SEQUENCE 527 AA; 55663 MW; A71E04E000A6D421 CRC64;

Query Match 79.6%; Score 156; DB 2; Length 527;  
Best Local Similarity 76.5%; Pred. No. 2e-05;  
Matches 26; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 34  
Db 192 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 225

## RESULT 15

AAR23813 PRELIMINARY; PRT; 527 AA.  
AC AAR23813;  
DT 02-MAR-2004 (TREMBlrel. 27, Created)  
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
DE DSXM.  
GN DSX.  
OS Musca domestica (House fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Muscoidea;  
OC Muscidae; Musca.  
OX NCBI\_TaxID=7370;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=14673649;  
RA Hediger M., Burghardt G., Siegenthaler C., Buser N.,  
RA Hilfinger-Kleiner D., Dubendorfer A., Bopp D.;  
RT "Sex determination in Drosophila melanogaster and Musca domestica  
converges at the level of the terminal regulator doublesex.";  
RL Dev. Genes Evol. 214:29-42(2004).  
DR EMBL; AY461854; AAR23813.1; -  
SQ SEQUENCE 527 AA; 55663 MW; A71E04E000A6D421 CRC64;

Query Match 79.6%; Score 156; DB 2; Length 527;  
Best Local Similarity 76.5%; Pred. No. 2e-05;  
Matches 26; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 34

Sun Nov 28 09:38:24 2004

us-09-936-702-2.rup

Page 6

D**b**

192 GCGGCGGCGGCGGCGGCGGCGGCGGCGGAGCGGNGGGC 225

Search completed: November 21, 2004, 13:42:23  
Job time : 21.4161 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:25:21 ; Search time 262.343 Seconds  
(without alignments)  
694.642 Million cell updates/sec

Title: US-09-936-702-3  
Perfect score: 2668  
Sequence: 1 MNRGVPRFHLVLQALLP.....GTRLEIKLVPRGSHHHHH 508

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2651.5	99.4	507	3	AAB00158
2	1064	39.9	295	2	AAR07605
3	1064	39.9	318	2	AAR07606
4	1064	39.9	394	2	AAY39825
5	1064	39.9	394	3	AAY88328
6	1064	39.9	416	3	AAB19509
7	1064	39.9	436	3	AAY51080
8	1064	39.9	458	3	AAY88329
9	1064	39.9	458	4	AAB81502
10	1064	39.9	458	4	AAG79087
11	1064	39.9	458	7	ADD25609
12	1064	39.9	458	7	ADDS5609
13	1064	39.9	458	8	ADBS7489
14	1064	39.9	458	8	ADP12485
15	1064	39.9	473	7	ADA44807
16	1064	39.9	474	3	AAY59170
17	1064	39.9	481	3	AAB19510
18	1064	39.9	481	3	AAY51081
19	1064	39.9	481	3	AAY59171
20	1064	39.9	519	2	AAR20152
21	1064	39.9	616	3	AAY51082
22	1064	39.9	616	3	AAY59172
23	1064	39.9	631	1	AAP93009
24	1064	39.9	631	3	AAB19508
25	1064	39.9	631	3	AAY51079
					AAY59169
					CD4-Ig fu

## ALIGNMENTS

26	1064	39.9	729	1	AAP93008	AAP93008 Genetic c
27	1064	39.9	729	3	AAB19507	AAB19507 CD4-IgG1
28	1064	39.9	729	3	AAY51078	AAY51078 Human fus
29	1064	39.9	729	3	AAY59168	AAY59168 CD4-Ig fu
30	1062	39.8	458	7	ADBS5841	ADBS5841 Human CD4
31	1059	39.7	398	2	AAR78673	AAR78673 CD4 domai
32	1059	39.7	398	2	AAR89450	AAR89450 CD4 D1-D4
33	1059	39.7	400	2	AAR06374	AAR06374 Truncated
34	1059	39.7	402	1	AAP91922	AAP91922 Sequence
35	1059	39.7	402	1	AAP94757	AAP94757 Sequence
36	1059	39.7	458	1	AAP81990	AAP81990 Clone pT4
37	1059	39.7	458	1	AAP91369	AAP91369 T4 protei
38	1059	39.7	458	2	AAR13491	AAR13491 Human CD4
39	1059	39.7	458	2	AAY39826	AAY39826 Soluble h
40	1059	39.7	458	3	AAB07769	AAB07769 DNA encod
41	1059	39.7	462	2	AAR27277	AAR27277 CD4:eta p
42	1059	39.7	462	2	AAR78677	AAR78677 T-cell re
43	1059	39.7	462	2	AAR89457	AAR89457 CD4:gamma
44	1059	39.7	462	2	AAW02214	AAW02214 CD4:Fc re
45	1059	39.7	462	2	AAW83142	AAW83142 Chimeric

RESULT 1  
AAB00158  
ID AAB00158 standard; protein; 507 AA.  
AC  
XX AAB00158;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE sCD4-SCFV(17b) HIV single chain antibody fusion protein.  
XX  
XX Fusion protein: HIV; human immunodeficiency virus; antibody; Fv; AIDS;  
XX acquired immune deficiency syndrome; neutralisation; infection;  
XX gene therapy; CD4; gp120; glycoprotein; resistance; vaccination;  
XX binding domain; single chain antibody; chimera; chimeric protein.  
OS Human immunodeficiency virus.  
OS Synthetic.  
PN WO200055207-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 16-MAR-2000; 2000MO-US006946.  
XX  
XX 16-MAR-1999; 99US-0124681P.  
XX  
XX (USSH ) US NAT INST OF HEALTH.  
XX  
XX Berger EA, Del Castillo CM;  
XX WPI; 2000-638183/61.  
XX N-PSDB; AAY54045.  
XX  
XX Novel neutralizing bispecific fusion proteins effective in viral such as  
XX HIV neutralization, comprises two different binding domains, inducing-  
XX binding domain and induced-binding domain functionally linked by linker.  
XX  
XX Claim 39; Page 46-47; 55pp; English.  
XX  
XX sCD4-SCFV(17b) is a neutralising bispecific fusion protein capable of  
XX binding to two sites of its target protein. The protein comprises a first  
XX binding domain capable of binding to an inducing site on the target  
XX protein, a second binding domain capable of forming neutralising complex  
XX with an induced epitope of the target protein and a linker connecting the  
XX binding domains. sCD4-SCFV(17b) comprises a soluble CD4 fragment  
XX (containing domains D1 and D2) fused to a single chain Fv portion of  
XX antibody 17b via a linker. sCD4-SCFV(17b), its variant, analogue or  
XX mimetic is used for inactivating gp120 protein of HIV, and for

CC neutralising HIV. It is also used for blocking and preventing the binding  
 CC of the viral or recombinant gp120 protein to soluble CD4 or lymphocyte  
 CC CD4 and for inhibiting HIV replication. The chimeric proteins is  
 CC therefore useful for treating HIV infection and also AIDS. It is are  
 CC particularly useful in the prevention of infection during or immediately  
 CC after HIV exposure (e.g., mother/infant transmission, post-exposure  
 CC prophylaxis, and as a topical inhibitor) and for providing long term  
 CC resistance to HIV infections and AIDS. Gene therapy is used to secrete  
 CC the bispecific protein at mucosal surfaces, such as the vaginal, rectal  
 CC or oral mucosa. The fusion proteins is highly potent, broadly cross-  
 CC reactive with neutralising antibody with high in vivo activity and no Fc-  
 CC mediated undesirable targeting properties. When the fusion protein is  
 CC substantially derived from human proteins, it has minimal immunogenicity  
 CC and toxicity in humans which is of great value in prevention of infection  
 CC during or immediately after HIV exposure

XX  
 XX Sequence 507 AA;

Query Match 99.4%; Score 2651.5; DB 3; Length 507;

Best Local Similarity 99.8%; Pred. No. 9e-128; Matches 507; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MNRGVPFRHLHLVQLALPRAATQKKVILGKGDVVELTCTASQKKSIOFHMNSNOIK 60  
 DB 1 MNRGVPFRHLHLVQLALPRAATQKKVILGKGDVVELTCTASQKKSIOFHMNSNOIK 60  
 QY 61 ILNGQSFLLTKGSPKNDRAADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120  
 DB 61 ILNGQSFLLTKGSPKNDRAADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120  
 QY 121 LVFELTANSDTHLLQGSLLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFELTANSDTHLLQGSLLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWICTVILQNOQKVEFKIDIVLAFOKASGGGSGGGSGGGSGGGSGGGSGGGSGGG 240  
 DB 181 TWICTVILQNOQKVEFKIDIVLAFOKASGGGSGGGSGGGSGGGSGGGSGGGSGGG 240  
 QY 241 GGSQVOLLGSAEYKPKSSVYKSCASGDTFIRYSFTWVQAQGGLEMMGRITITLDV 300  
 DB 241 GGSQVOLLGSAEYKPKSSVYKSCASGDTFIRYSFTWVQAQGGLEMMGRITITLDV 300  
 QY 301 AHYAPHLQGRVITITADKSTSTVYLELNLRSDPTAVVFCAGVYGEADEGEYDNNKFLKH 360  
 DB 301 AHYAPHLQGRVITITADKSTSTVYLELNLRSDPTAVVFCAGVYGEADEGEYDNNKFLKH 360  
 QY 361 WGGTILVTITSGGGSGGGSGGGSELTLTQSPATLVSPPERATLSCRASESVSDDLA 420  
 DB 361 WGGTILVTITSGGGSGGGSGGGSELTLTQSPATLVSPPERATLSCRASESVSDDLA 420  
 QY 421 WYQCKPQAPRLLLYGASTRATGVPARFSGSGAEFTLTSSIQSDFPAYVYCOQNNW 480  
 DB 421 WYQCKPQAPRLLLYGASTRATGVPARFSGSGAEFTLTSSIQSDFPAYVYCOQNNW 480  
 QY 481 PPRYTFQGTLEIKLIVPRSGHHNNH 508  
 DB 481 PPRYTFQGTLEIKLIVPRSGHHNNH 508  
 QY 481 PPRYTFQGTLEIKLIVPRSGHHNNH 508  
 DB 481 PPRYTFQGTLEIKLIVPRSGHHNNH 508

RESULT 2  
 AAR07605  
 ID AAR07605 standard; protein; 295 AA.

AC AAR07605;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 09-JAN-2003 (revised)  
 DT 06-FEB-1991 (first entry)

DE Plasmid T4/LFA-3/AD gene product with C-terminal encoding  
 DE phosphatidylinositol (PI) linkage signaling sequence.  
 XX  
 XX Plasma membrane binding affinity; micelle.

XX Unidentified.

XX Key Location/Qualifiers

FT Protein 1..266 /label= T4/LFA-3/AD gene product

FT Peptide 267..295 /label= PI signal peptide

XX MO9012099-A.

XX 18-OCT-1990.

XX 10-APR-1989; 89US-00335688.

XX 10-APR-1989; 89US-00335688.

XX (BIOJ ) BIOGEN INC.

XX Wallner BP;

XX WPI; 1990-334849/44.

XX N-PSDB; AAQ06404.

XX Phosphatidylinositol linkage signalling DNA sequence - derived from  
 PT lymphocyte function-associated antigen 3, used for prodn. of chimeric  
 PT proteins.

XX Disclosure; Fig 4; 53pp; English.

XX The signal sequence is attached downstream of the plasmid sequence  
 CC encoding a secreted protein, which will then produce proteins covalently  
 CC anchored to the cell surface in which they were produced. This can give  
 CC rise to plasma membrane binding, enhanced purifiability, micelle  
 CC formation etc. especially useful in the production of chimeric targeted  
 CC drugs, to produce micellar or liposomal delivery systems or in enhanced  
 CC purification and screening of cells, proteins or DNA libraries. (Updated  
 CC on 09-JAN-2003 to add missing OS field.) (updated on 25-MAR-2003 to  
 CC correct PA field.)

XX Sequence 295 AA;

Query Match 39.9%; Score 1064; DB 2; Length 295;

Best Local Similarity 100.0%; Pred. No. 6.2e-47; Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLHLVQLALPRAATQKKVILGKGDVVELTCTASQKKSIOFHMNSNOIK 60  
 DB 1 MNRGVPFRHLHLVQLALPRAATQKKVILGKGDVVELTCTASQKKSIOFHMNSNOIK 60  
 QY 61 ILNGQSFLLTKGSPKNDRAADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120  
 DB 61 ILNGQSFLLTKGSPKNDRAADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120  
 QY 121 LVFELTANSDTHLLQGSLLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFELTANSDTHLLQGSLLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWICTVILQNOQKVEFKIDIVLAFOKAS 208  
 DB 181 TWICTVILQNOQKVEFKIDIVLAFOKAS 208

RESULT 3  
 AAR07606  
 ID AAR07606 standard; protein; 318 AA.

AC AAR07606;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 09-JAN-2003 (revised)  
 DT 06-FEB-1991 (first entry)



DE Plasmid T4/LFA-3/2 gene product with C-terminal encoding  
DE phosphatidylinositol (PI) linkage signaling sequence.  
KW Plasma membrane binding affinity; micelle.  
XX Unidentified.  
OS  
XX  
XX  
FH Key Location/Qualifiers  
FT Protein 1..267  
FT /label= LFA-3 gene product  
FT Peptide 267..318  
FT /label= PI signal peptide  
XX  
XX WO9012099-A.  
XX  
XX 18-OCT-1990.  
XX  
XX 10-APR-1989; 89US-00335688.  
XX  
XX 10-APR-1989; 89US-00335688.  
XX  
XX (BIOJ ) BIOGEN INC.  
XX  
XX Wallner BP;  
XX  
XX WPI, 1990-334849/44.  
XX  
XX N-PSDB; AAQ06405.  
XX  
XX Phosphatidylinositol linkage signalling DNA sequence - derived from  
XX lymphocyte function-associated antigen 3, used for prodn. of chimeric  
XX proteins.  
XX  
XX Disclosure; Fig 5; 53pp; English.  
XX  
XX The signal sequence is attached downstream of the Plasmid sequence  
XX encoding a secreted protein, which will then produce proteins covalently  
XX anchored to the cell surface in which they were produced. This can give  
XX rise to plasma membrane binding, enhanced purifiability, micelle  
XX formation etc. especially useful in the production of chimeric targeted  
XX drugs, to produce micellar or liposomal delivery systems or in enhanced  
XX purification and screening of cells, proteins or DNA libraries. (Updated  
XX on 09-FAN-2003 to add missing OS field.) (Updated on 25-MAR-2003 to  
XX correct PA field.)  
XX  
SQ Sequence 318 AA;  
Query Match 39.9%; Score 1064; DB 2; Length 318;  
Best Local Similarity 100.0%; Pred. No. 6.6e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNRGVPRHLLVLOLALPPATQGGKVVLGKGDVTELTCTASOKSIOFHWKNSNOIK 60  
DB 1 MNRGVPRHLLVLOLALPPATQGGKVVLGKGDVTELTCTASOKSIOFHWKNSNOIK 60  
QY ILNGGSFLTKGPSKLNDRADSRSLMDQGNFPLIKKLTIEDSPTYICEVEDQKEVQL 120  
DB 61 ILNGGSFLTKGPSKLNDRADSRSLMDQGNFPLIKKLTIEDSPTYICEVEDQKEVQL 120  
QY 121 LVFGILTANSDTHLLQGOSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGILTANSDTHLLQGOSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLOKQKVEFKIDIVVLAFOKAS 208  
DB 181 TWTCTVLOKQKVEFKIDIVVLAFOKAS 208  
RESULT 4  
AAV39825  
ID AAV39825 standard; protein; 394 AA.  
XX AAV39825;  
XX

DT 03-DEC-1999 (first entry)  
XX  
XX Soluble human T4 protein.  
XX  
XX Soluble T4 protein; sT4; human; HIV; binding inhibitor; T4+ cell; AIDS;  
KW vaccine; immunisation; therapy.  
XX  
XX Homo sapiens.  
XX  
XX US5958678-A.  
XX  
XX 28-SEP-1999.  
XX  
XX 12-DEC-1994; 94US-00354452.  
XX  
XX 21-AUG-1986; 86US-00898587.  
XX  
XX 11-JUN-1991; 91US-00713564.  
XX  
XX 06-JUL-1992; 92US-00909021.  
XX  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
XX  
XX McDougal JS, Weiss R, Axel R, Littman DR, Maddon PJ, Chess L;  
XX  
XX WPI, 1999-561025/47.  
XX  
XX N-PSDB; AAZ20694.  
XX  
XX Human T4 protein inhibits HIV binding to T4 cells, useful for treating  
XX AIDS.  
XX  
XX Disclosure; Col 13-16; 58pp; English.  
XX  
XX This sequence represents the soluble human T4 protein of the invention.  
XX The soluble human T4 protein blocks the binding of HIV to T4+ cells and  
XX is therefore useful for the treatment of AIDS. Monoclonal antibodies  
XX against the T4 protein may be used as vaccines for immunising subjects  
XX against AIDS  
XX  
SQ Sequence 394 AA;  
Query Match 39.9%; Score 1064; DB 2; Length 394;  
Best Local Similarity 100.0%; Pred. No. 7.9e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNRGVPRHLLVLOLALPPATQGGKVVLGKGDVTELTCTASOKSIOFHWKNSNOIK 60  
DB 1 MNRGVPRHLLVLOLALPPATQGGKVVLGKGDVTELTCTASOKSIOFHWKNSNOIK 60  
QY ILNGGSFLTKGPSKLNDRADSRSLMDQGNFPLIKKLTIEDSPTYICEVEDQKEVQL 120  
DB 61 ILNGGSFLTKGPSKLNDRADSRSLMDQGNFPLIKKLTIEDSPTYICEVEDQKEVQL 120  
QY 121 LVFGILTANSDTHLLQGOSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGILTANSDTHLLQGOSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLOKQKVEFKIDIVVLAFOKAS 208  
DB 181 TWTCTVLOKQKVEFKIDIVVLAFOKAS 208  
RESULT 5  
AAV88328  
ID AAV88328 standard; protein; 394 AA.  
XX AAV88328;  
XX  
XX 14-JUL-2000 (first entry)  
XX  
XX T4 glycoprotein amino acid sequence.  
XX  
XX sT4; glycoprotein; human immunodeficiency virus; HIV; block binding;  
KW AIDS; treatment; inhibit; cell to cell spread; infection; fusion.  
XX

OS Mammalia.  
 XX US5126433-A.  
 PN 30-JUN-1992.  
 XX 23-OCT-1987; 87US-00114244.  
 PF 21-AUG-1986; 86US-00898587.  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 PA Maddon PJ, Chess L, Axel R, Weiss R, Littman DR, McDougal JS;  
 PI WPI; 2000-348913/30.  
 DR N-PSDB; AAA10906.  
 XX Soluble T4 glycoprotein useful for prevention and treatment of acquired  
 PT immunodeficiency syndrome and for screening inhibitors of human  
 PT immunodeficiency viral binding.  
 XX  
 PS Disclosure; Col 11-16; 64pp; English.  
 XX  
 CC This sequence represents the full length amino acid sequence of  
 CC glycosylated sT4 glycoprotein. Human immunodeficiency virus (HIV) uses  
 CC sT4 as a target receptor on T cells. The invention relates to  
 CC glycosylated sT4 which functions by blocking the binding of HIV to T4  
 CC target cells, and can be used for the prophylaxis and treatment of AIDS  
 CC patients. Administration of sT4 effectively inhibits the cell to cell  
 CC spreading of HIV infection and also the fusion of HIV-infected T4 cells  
 CC and non-infected T4 cells. The administration of T4 alleviates several  
 CC symptoms associated with AIDS, and prevents the occurrence of new  
 CC pathological changes. The sT4 glycoprotein is useful for the prophylaxis  
 CC and treatment of patients with AIDS. It is also useful as a reagent to  
 CC identify natural, synthetic or recombinant molecules which act as  
 CC therapeutic agents or inhibitors of T4+ cell interactions and in  
 CC diagnostic assays for detection of T4 proteins or molecules  
 XX  
 SQ Sequence 394 AA;  
 Query Match 39.9%; Score 1064; DB 3; Length 394;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MNRGVPFRHLVLTQALLPAAATQGGKVVLGKGDVVELTCTASQKKSIOFHMKNNSQIK 60  
 DB 1 MNRGVPFRHLVLTQALLPAAATQGGKVVLGKGDVVELTCTASQKKSIOFHMKNNSQIK 60  
 QY 61 ILNGQGSFLTKGSPSKLNDRAISRSLWDQGNFPLIIKNLKIEDSDTYICEVEDQKEVQL 120  
 DB 61 ILNGQGSFLTKGSPSKLNDRAISRSLWDQGNFPLIIKNLKIEDSDTYICEVEDQKEVQL 120  
 QY 121 LVFGLTANSDTHLLQGGSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDTHLLQGGSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWTCTVLOQKQKVEFKIDIVLAFOKAS 208  
 DB 181 TWTCTVLOQKQKVEFKIDIVLAFOKAS 208  
 RESULT 6  
 AAB19509  
 ID AAB19509 standard; protein; 416 AA.  
 XX AAB19509;  
 AC  
 XX  
 DT 09-JAN-2001 (first entry)  
 XX  
 DE CD4-IgM fusion protein CH4Mmu.  
 XX  
 KM CD4; IGM; human; CD4Mmu; fusion protein; immunoglobulin; HIV; SIV; gp120;  
 therapy; diagnosis.

XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT Protein  
 FT Protein  
 FT Protein  
 XX US6117656-A.  
 XX  
 XX 12-SEP-2000.  
 PD 07-JUN-1995; 95US-00479353.  
 PF  
 XX 22-JAN-1988; 88US-00147351.  
 PR 23-JAN-1989; 89US-00299596.  
 PR 09-JUN-1992; 92US-00896781.  
 PR 12-APR-1993; 93US-00057952.  
 PR 04-FEB-1994; 94US-00191708.  
 XX  
 PA (GEHO ) GEN HOSPITAL CORP.  
 XX  
 PI Seed B;  
 XX  
 DR WPI; 2000-586558/55.  
 DR N-PSDB; AAA50662.  
 XX  
 PT CD4-immunoglobulin fusion proteins, useful for targeting gp120 of HIV or  
 PT SIV.  
 PS Example 1; Col 41-50; 39pp; English.  
 XX  
 CC The present sequence is that of fusion protein CD4Mmu comprising the  
 CC extracellular portion of CD4, which binds to HIV gp120, linked at its C-  
 CC terminus to the human Igm heavy chain. To obtain the fusion protein, DNA  
 CC encoding CD4 was linked to Igm DNA at the MscI site upstream of the CH1  
 CC region (see AAA50662). Fusion protein CD4Mmu and a nucleic acid encoding  
 CC it are claimed. Also claimed are a vector comprising the nucleic acid,  
 CC and a method of producing the fusion protein in secreted form using a  
 CC transformed host cell. The fusion protein may further comprise a  
 CC therapeutic agent, radiolabel or NMR imaging agent. The fusion protein  
 CC can be administered to an animal (including humans) for treatment of HIV  
 CC or SIV infection, and can also be used in assays for HIV or SIV, imaging  
 CC and tissue stains. Igm fusion proteins such as CD4Mmu provide complement-  
 CC mediated immunity  
 XX  
 SQ Sequence 416 AA;  
 Query Match 39.9%; Score 1064; DB 3; Length 416;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MNRGVPFRHLVLTQALLPAAATQGGKVVLGKGDVVELTCTASQKKSIOFHMKNNSQIK 60  
 DB 1 MNRGVPFRHLVLTQALLPAAATQGGKVVLGKGDVVELTCTASQKKSIOFHMKNNSQIK 60  
 QY 61 ILNGQGSFLTKGSPSKLNDRAISRSLWDQGNFPLIIKNLKIEDSDTYICEVEDQKEVQL 120  
 DB 61 ILNGQGSFLTKGSPSKLNDRAISRSLWDQGNFPLIIKNLKIEDSDTYICEVEDQKEVQL 120  
 QY 121 LVFGLTANSDTHLLQGGSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDTHLLQGGSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWTCTVLOQKQKVEFKIDIVLAFOKAS 208  
 DB 181 TWTCTVLOQKQKVEFKIDIVLAFOKAS 208  
 RESULT 7  
 AA51080  
 ID AA51080 standard; protein; 436 AA.

XX AC AAY51080; (first entry)  
XX DT 23-MAR-2000  
XX DE Human fusion protein CD4Mg.  
XX DE Human fusion protein CD4Mg.  
XX DE Fusion protein; human; CD4; IGM; immunoglobulin; gp120;  
XX DE anti-human immunodeficiency virus; CD4Mg.  
XX OS Homo sapiens.  
XX OS Synthetic.  
XX PN US6004781-A.  
XX PD 21-DEC-1999.  
XX PF 04-FEB-1994; 94US-00191708.  
XX PR 22-JAN-1988; 88US-00147351.  
XX PR 23-JAN-1989; 89US-00299596.  
XX PR 09-JUN-1992; 92US-0086781.  
XX PR 12-APR-1993; 93US-00057952.  
XX PA (GHEO) GEN HOSPITAL CORP.  
XX PI Seed B;  
XX DR WPI; 2000-085792/07.  
XX DR N-PSDB; AA244063.  
XX PS Fusion protein useful for the treatment of human immunodeficiency virus.  
XX PT Example 1; Col 41-50; 39pp; English.  
XX PS This invention describes a novel nucleic acid (I) encoding a fusion  
XX CC protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)  
XX CC and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light  
XX CC chain (III). The products of the invention have anti-human  
XX CC immunodeficiency virus (HIV) activity and are capable of binding to  
XX CC gp120. The fusion protein is useful for treating human immunodeficiency  
XX CC virus (HIV) or simian immunodeficiency virus (SIV). This sequence  
XX CC represents the fusion protein CD4Mg which is constructed from CD4 linked  
XX CC to human IGM upstream of the CH1 region  
XX SQ Sequence 436 AA;  
Query Match 39.9%; Score 1064; DB 3; Length 436;  
Best Local Similarity 100.0%; Pred. No. 8.7e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX XX 14-JUL-2000 (first entry)  
XX DT T4 glycoprotein amino acid sequence.  
XX DE T4 glycoprotein; human immunodeficiency virus; HIV; block binding;  
XX DE sT4; glycoprotein; human immunodeficiency virus; HIV; block binding;  
XX DE AIDS; treatment; inhibit; cell to cell spread; infection; fusion.  
XX OS Mammalia.  
XX PN US5126433-A.  
XX PD 30-JUN-1992.  
XX PF 23-OCT-1987; 87US-00114244.  
XX PR 21-AUG-1986; 86US-00898587.  
XX PA (UYCO) UNIV COLUMBIA NEW YORK.  
XX PI Maddon PJ, Chess L, Axel R, Weiss R, Littman DR, McDougal JS;  
XX DR WPI; 2000-348913/30.  
XX PT Soluble T4 glycoprotein useful for prevention and treatment of acquired  
XX PT immunodeficiency syndrome and for screening inhibitors of human  
XX PS immunodeficiency viral binding.  
XX PS Example; Fig 6; 64pp; English.  
XX CC This sequence represents the amino acid sequence of glycosylated sT4  
XX CC glycoprotein. Human immunodeficiency virus (HIV) uses sT4 as a target  
XX CC receptor on T cells. The invention relates to glycosylated sT4 which  
XX CC functions by blocking the binding of HIV to T4 target cells, and can be  
XX CC used for the prophylaxis and treatment of AIDS patients. Administration  
XX CC of sT4 effectively inhibits the cell to cell spreading of HIV infection  
XX CC and also the fusion of HIV-infected T4 cells and non-infected T4 cells.  
XX CC The administration of T4 alleviates several symptoms associated with  
XX CC AIDS, and prevents the occurrence of new pathological changes. The sT4  
XX CC glycoprotein is useful for the prophylaxis and treatment of patients with  
XX CC AIDS. It is also useful as a reagent to identify natural, synthetic or  
XX CC recombinant molecules which act as therapeutic agents or inhibitors of  
XX CC T4 cell interactions and in diagnostic assays for detection T4 proteins  
XX CC or molecules  
XX SQ Sequence 458 AA;  
Query Match 39.9%; Score 1064; DB 3; Length 458;  
Best Local Similarity 100.0%; Pred. No. 9e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 18-JUN-2001 (first entry)  
 DT Human CD4 protein.  
 XX  
 DE Human CD4 protein.  
 XX  
 KW Human; CD4; CD4 fusion protein; oligomerisation;  
 KW receptor-ligand interaction inhibition; surface plasmon resonance; SPR;  
 KW T cell receptor binding; MHC binding; carcinoma; autoimmune disease;  
 KW multiple sclerosis; human immunodeficiency virus; HIV; diabetes;  
 KW rheumatoid arthritis; immune disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT Peptide 1..25  
 FT /label= Signal\_peptide  
 FT Protein 26..458  
 FT /label= Human\_CD4  
 XX  
 PN WO200122084-A2.  
 XX  
 PD 29-MAR-2001.  
 XX  
 PF 18-SEP-2000; 2000WO-GB003579.  
 XX  
 PR 21-SEP-1999; 99GB-00022352.  
 XX  
 PA (AVID-) AVIDEX LTD.  
 XX  
 PI Jakobsen BK;  
 XX  
 DR WPI; 2001-273470/28.  
 DR N-PSDB; AAF82582.  
 XX  
 PT Sequential screening of candidate compounds library for those which  
 PT inhibit binding of low affinity receptor-ligand interaction having fast  
 PT binding kinetics, using interfacial optical assay.  
 XX  
 PS Disclosure; Fig 13; 91pp; English.  
 XX  
 CC The present sequence is human CD4. Human CD4 extracellular domains 1 and  
 CC 2 were used in the construction of CD4 oligomerisation fusion proteins.  
 CC The fusion proteins contain an oligomerisation domain that enables the  
 CC proteins to bind to one another to form oligomers. The oligomers may be  
 CC used in an invention relating to a method for screening for compounds  
 CC with the ability to inhibit a low affinity receptor-ligand interaction.  
 CC The method uses an interfacial optical assay, such as surface plasmon  
 CC resonance (SPR). The method is useful for screening candidate compounds  
 CC for the ability to inhibit interaction between MHC/peptide complex and T  
 CC cell receptor, and MHC/peptide complex and CD8 or CD4 co-receptor. The  
 CC compounds identified by the above methods which interfere with T cell  
 CC receptor binding to a particular HLA type molecule are useful as immune  
 CC inhibitors for treating carcinomas, autoimmune diseases such as immune  
 CC sclerosis, human immunodeficiency virus (HIV) infection, rheumatoid  
 CC arthritis, Hashimoto's disease, insulin dependent diabetes, Good  
 CC pasture's syndrome, uveitis, psoriasis and graft rejection  
 CC  
 XX  
 SQ Sequence 458 AA;  
 Query Match 39.9%; Score 1064; DB 4; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 9e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MNRGVPFRHLLVQLALPAATQGGKVVLGKGGDTVELTCTASOKKSIOFHMKNSNOIK 60  
 DB 1 MNRGVPFRHLLVQLALPAATQGGKVVLGKGGDTVELTCTASOKKSIOFHMKNSNOIK 60  
 QY 1 ILNGQGSFLTTPGSKLNDRADSRSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVQL 120  
 DB 61 ILNGQGSFLTTPGSKLNDRADSRSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVQL 120  
 QY 1 ILNGQGSFLTTPGSKLNDRADSRSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVQL 120  
 DB 61 ILNGQGSFLTTPGSKLNDRADSRSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVQL 120  
 QY 121 LVFGLTANSSTHLLQGQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 DB 121 LVFGLTANSSTHLLQGQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180

DB 121 LVFGLTANSSTHLLQGQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 QY 181 TWTCVTLQNGQKVEFKIDIVLAFOKAS 208  
 DB 181 TWTCVTLQNGQKVEFKIDIVLAFOKAS 208  
 RESULT 10  
 ID AAG79087  
 AC AAG79087 standard; protein; 458 AA.  
 XX  
 AC AAG79087;  
 XX  
 DT 10-DEC-2001 (first entry)  
 XX  
 DE Amino acid sequence of human CD4 protein.  
 XX  
 KW Human; receptor; DC-SIGN; dendritic cell; T lymphocyte; HIV; gp120;  
 KW C-type lectin; ICAM3; HIV entry; T cell; macrophage; HIV infection; CD4.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164752-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 28-FEB-2001; 2001WO-US006322.  
 XX  
 PR 02-MAR-2000; 2000US-00517605.  
 XX  
 PA (UYNY ) UNIV NEW YORK STATE.  
 PA (UYNI-) UNIV NIJMEGEN.  
 XX  
 PI Littman DR, Kwon D, Van Kooyk Y, Geijtenbeek T;  
 XX  
 DR WPI; 2001-602565/68.  
 XX  
 PT An antibody for the treatment or prevention of HIV-infection comprises a  
 PT gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding of  
 PT DC-SIGN due to concomitant conformational change.  
 XX  
 PS Disclosure; Page 115-116; 131pp; English.  
 XX  
 CC The specification describes an antibody which is specific for an  
 CC antigenic fragment of gp120. This antigenic fragment binds to DC-SIGN or  
 CC is exposed upon gp120 binding of DC-SIGN due to concomitant  
 CC conformational change. DC-SIGN is a receptor that is specifically  
 CC expressed on dendritic cells and facilitates infection of T lymphocytes  
 CC with HIV. DC-SIGN is identical to a HIV-1 gp120-binding C-type lectin. DC  
 CC -SIGN binds ICAM-3 (which is expressed constitutively on T lymphocytes)  
 CC with high affinity. The antibody of the invention inhibits the trans  
 CC enhancement of HIV entry into a T cell or macrophage facilitated by  
 CC dendritic cells. The antibody is useful to treat or prevent HIV  
 CC infection. The present sequence represents a human CD4 protein  
 CC  
 XX  
 SQ Sequence 458 AA;  
 Query Match 39.9%; Score 1064; DB 4; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 9e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MNRGVPFRHLLVQLALPAATQGGKVVLGKGGDTVELTCTASOKKSIOFHMKNSNOIK 60  
 DB 1 MNRGVPFRHLLVQLALPAATQGGKVVLGKGGDTVELTCTASOKKSIOFHMKNSNOIK 60  
 QY 1 ILNGQGSFLTTPGSKLNDRADSRSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVQL 120  
 DB 61 ILNGQGSFLTTPGSKLNDRADSRSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVQL 120  
 QY 1 ILNGQGSFLTTPGSKLNDRADSRSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVQL 120  
 DB 61 ILNGQGSFLTTPGSKLNDRADSRSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVQL 120  
 QY 121 LVFGLTANSSTHLLQGQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 DB 121 LVFGLTANSSTHLLQGQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180

QY 181 TWTCVTLQNKQKVEFKIDIVLAFQKAS 208  
 |||||  
 DB 181 TWTCVTLQNKQKVEFKIDIVLAFQKAS 208

RESULT 11  
 ADD25609  
 ID ADD25609 standard; protein; 458 AA.  
 XX  
 AC ADD25609;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Binding domain-immunoglobulin fusion protein-associated protein #82.  
 XX  
 KW Binding domain; immunoglobulin; fusion protein; cytostatic;  
 KW antitachytic; immunosuppressive; antidiabetic; antithyroid;  
 KW neuroprotective; hinge region; immunoglobulin heavy chain;  
 KW CH2 constant region; CH3 constant region; IgG1;  
 KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;  
 KW rheumatoid arthritis; myasthenia gravis; Grave's disease;  
 KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.  
 XX  
 OS Unidentified.  
 PN US2003118592-A1.  
 XX  
 PD 26-JUN-2003.  
 XX  
 PF 25-JUL-2002; 2002US-00207655.  
 PR 17-JAN-2001; 2001US-0367358P.  
 PR 17-JAN-2002; 2002US-00053530.  
 PR 03-JUN-2002; 2002US-0385691P.  
 PA (GENE-) GENE-CRAFT INC.  
 XX  
 PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;  
 XX  
 DR WPI; 2003-801317/75.  
 XX  
 PT New binding domain-immunoglobulin fusion protein, useful for treating a  
 PT subject having or suspected of having a malignant condition or a B-cell  
 PT disorder, e.g., melanoma, Grave's disease or autoimmune disease.  
 XX  
 PS Disclosure; SEQ ID NO 170; 157bp; English.  
 XX  
 XX The invention relates to a binding domain-immunoglobulin fusion protein  
 XX comprising a binding domain polypeptide that is fused to an  
 XX immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain  
 XX CH2 constant region polypeptide that is fused to the hinge region  
 XX polypeptide, and an immunoglobulin-heavy chain CH3 constant region  
 XX polypeptide that is fused to the CH2 constant region polypeptide. The  
 XX hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin  
 XX hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge  
 XX region polypeptide, derived from (a) having 3 or more cysteine residues;  
 XX where the mutated human IgG1 immunoglobulin hinge region polypeptide  
 XX contains 2 cysteine residues, where the first cysteine is not mutated; a  
 XX mutated human IgG1 immunoglobulin hinge region polypeptide, derived from  
 XX (a) having 3 or more cysteine residues, where the mutated human IgG1  
 XX immunoglobulin hinge region polypeptide contains no more than one  
 XX cysteine residue; and a mutated human IgG1 immunoglobulin hinge region  
 XX polypeptide, derived from (a) having 3 or more cysteine residues; where  
 XX the mutated human IgG1 immunoglobulin hinge region polypeptide contains  
 XX no cysteine residues. The binding domain-immunoglobulin fusion protein is  
 XX capable of at least one immunological activity comprising antibody  
 XX dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The  
 XX binding domain polypeptide is capable of specifically binding to an  
 XX antigen. Also included are an isolated polynucleotide encoding the  
 XX binding domain-immunoglobulin fusion protein, a recombinant expression  
 XX construct comprising the polynucleotide (operably linked to a promoter),  
 XX a host cell transformed or transfected with a recombinant expression

CC construct, producing the binding domain-immunoglobulin fusion protein, a  
 CC pharmaceutical composition comprising the binding domain-immunoglobulin  
 CC fusion protein or polynucleotide and a carrier, and treating a subject  
 CC having or suspected of having a malignant condition or a B-cell disorder.  
 CC The binding domain-immunoglobulin fusion protein is useful for treating a  
 CC subject having or suspected of having a malignant condition or a B-cell  
 CC disorder, e.g., melanoma, carcinoma or sarcoma, rheumatoid arthritis,  
 CC myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple  
 CC sclerosis or autoimmune disease. The present sequence is a binding domain  
 CC -immunoglobulin fusion protein-associated protein sequence. Note: The  
 CC sequence data for this patent formed part of the printed specification  
 CC and is also available in electronic format directly from USPTO at  
 CC [segdata.uspto.gov/sequence.html?docID=20030118592](http://segdata.uspto.gov/sequence.html?docID=20030118592). The authors have not  
 CC identified the sequences in the printed specification by their SEQ ID  
 CC number therefore none of the sequences can be explicitly identified.  
 XX  
 XX SQ Sequence 458 AA;  
 XX  
 QY Query Match 39.9%; Score 1064; DB 7; Length 458;  
 DB Best Local Similarity 100.0%; Pred. NO. 9e-47;  
 DB Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MNRGVPRHLLVLTALLPATQKVVLGKGDVTELTCTASQKSIQFMKNSQIK 60  
 DB 1 MNRGVPRHLLVLTALLPATQKVVLGKGDVTELTCTASQKSIQFMKNSQIK 60  
 QY 61 ILNGQSFVTGPKSLNDRADSRSLMDQGNPLIINKLIEDSDTYICEVEDQKEEYOL 120  
 DB 61 ILNGQSFVTGPKSLNDRADSRSLMDQGNPLIINKLIEDSDTYICEVEDQKEEYOL 120  
 QY 121 LVFGLTANSDFHLLQGGSLTTLTSPSSSSVOCRSRGNIGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDFHLLQGGSLTTLTSPSSSSVOCRSRGNIGGKTLVSQLELDQSG 180  
 QY 181 TWTCVTLQNKQKVEFKIDIVLAFQKAS 208  
 DB 181 TWTCVTLQNKQKVEFKIDIVLAFQKAS 208

RESULT 12  
 ADE57489  
 ID ADE57489 standard; protein; 458 AA.  
 XX  
 AC ADE57489;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human Protein P01730, SEQ ID NO 3351.  
 XX  
 KW Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.  
 OS Homo sapiens.  
 XX  
 PN WO2003016475-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 14-AUG-2002; 2002WO-US025765.  
 PR 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 PA (GENO) GEN HOSPITAL CORP.  
 PA (FAR) BAYER AG.  
 XX  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX  
 DR WPI; 2003-268312/26.  
 DR GENBANK; P01730.

PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 458 AA:

Query Match 39.9%; Score 1064; DB 7; Length 458;  
Best Local Similarity 100.0%; Pred. No. 9e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRRHLLVQLALPPAATQGGKVVGGKGDVVELTCTASOKKSIOFHMKNSNQIK 60  
DB 1 MNRGVPRRHLLVQLALPPAATQGGKVVGGKGDVVELTCTASOKKSIOFHMKNSNQIK 60  
QY 61 ILNGQSFLLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILNGQSFLLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180  
DB 121 LVFGLTANSDTHLLQGGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180  
QY 181 TWTCTVLOKQKVEFKIDIVLAFOKAS 208  
DB 181 TWTCTVLOKQKVEFKIDIVLAFOKAS 208

RESULT 13

ADP12495  
ID ADP12495 standard; protein; 458 AA.

XX ADP12495;

DT 12-AUG-2004 (first entry)

DE Protein encoded by mRNA of the invention #105.

XX transplant rejection; immune system; rheumatoid arthritis; lupus;  
KM inflammatory bowel disease; multiple sclerosis; HIV; AIDS.

OS Homo sapiens.

XX MO2004042346-A2.  
XX PN

PD 21-MAY-2004.

XX 24-APR-2003; 2003MO-US012946.

XX 24-APR-2002; 2002US-00131831.

PR 20-DEC-2002; 2002US-00325899.

XX (EXPR-) EXPRESSION DIAGNOSTICS INC.

PI Wohlgenuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;  
PI Rosenberg S;

DR WPI; 2004-400724/37.

PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,  
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant  
PT rejection, in an individual, comprises detecting the expression level of  
PT the genes.

XX Claim 65; SEQ ID NO 2504; 1762pp; English.

CC The present invention relates to diagnosing or monitoring transplant  
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual  
CC comprising detecting the expression level of one or more genes. The  
CC methods, system and kits are useful in diagnosing or monitoring  
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic  
CC islet, lung, bone marrow or stem cell transplant rejection,  
CC xenotransplant rejection or mechanical organ replacement rejection, in an  
CC individual. The method is also useful in assessing the immune status of  
CC an individual. The methods are also useful in diagnosing and monitoring  
CC diseases that involve the immune system, e.g. rheumatoid arthritis,  
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or  
CC viral, bacterial or fungal infection. The present sequence represents a  
CC protein that is encoded by the mRNA of the invention.

XX Sequence 458 AA:

Query Match 39.9%; Score 1064; DB 8; Length 458;  
Best Local Similarity 100.0%; Pred. No. 9e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRRHLLVQLALPPAATQGGKVVGGKGDVVELTCTASOKKSIOFHMKNSNQIK 60  
DB 1 MNRGVPRRHLLVQLALPPAATQGGKVVGGKGDVVELTCTASOKKSIOFHMKNSNQIK 60  
QY 61 ILNGQSFLLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILNGQSFLLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180  
DB 121 LVFGLTANSDTHLLQGGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180  
QY 181 TWTCTVLOKQKVEFKIDIVLAFOKAS 208  
DB 181 TWTCTVLOKQKVEFKIDIVLAFOKAS 208

RESULT 14

ADA44807  
ID ADA44807 standard; protein; 473 AA.

XX ADA44807;

DT 04-DEC-2003 (first entry)

DE CD4/TCR CD3epsilon chain chimeric protein CD4epsilon105, SEQ ID NO:2.

XX HIV-1 infection; human immunodeficiency virus-1; CD4+ cell; chimeric CD4;  
KM endoplasmic reticulum; ER retention; envelope protein gp160;  
KM T cell receptor CD3epsilon chain; C-terminal domain; CD4epsilon105;  
KM gene therapy; human; receptor.

OS Chimeric.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Protein 1..458  
FT /label= CD4  
FT Region 459..473  
FT /note= "Part of the C-terminal domain of the T cell  
FT receptor CD3epsilon chain"  
XX  
XX WO2003076468-A1.  
XX  
XX 18-SEP-2003.  
XX  
XX 14-MAR-2003; 2003WO-ES000120.  
XX  
XX 14-MAR-2002; 2002ES-00000616.  
XX  
XX (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
XX  
XX Alarcon Sanchez BJ, San Jose Martinez ME, Zaldivar Notario I;  
XX Gomez Buendia M;  
XX WPI; 2003-779059/73.  
XX  
XX N-PSDB; ADA44806.  
XX  
XX Composition for treating or preventing human immune deficiency virus,  
XX comprises CD4 chimeric protein having a protective effect in trans, or  
XX related nucleic acid.  
XX  
XX Claim 5, Page 33-35; 43pp; Spanish.  
XX  
XX The invention relates to a composition for the treatment or prevention of  
XX human immunodeficiency virus-1 (HIV-1) infection. The composition  
XX comprises CD4+ cells that have been transduced with a vector that encodes  
XX a chimeric CD4 molecule which is capable of being retained in the  
XX endoplasmic reticulum (ER). The invention also encompasses the use of a  
XX soluble protein factor produced by CD4+ cells that have been transduced  
XX with a vector encoding a chimeric CD4 protein; and the use of an  
XX expression system encoding a chimeric CD4 protein. The ER-localised  
XX chimeric CD4 molecule binds to the HIV-1 envelope protein gp160,  
XX resulting in HIV-1 retention in the ER and thereby preventing viral  
XX replication. In a specific embodiment, the chimeric CD4 molecule  
XX comprises CD4 fused to 15 amino acids of the C-terminal domain of the T  
XX cell receptor CD3epsilon chain; this chimeric CD4 molecule is designated  
XX CD4epsilon1015 (ADA44807). A known chimeric CD4 of similar structure but  
XX containing only 10 amino acids from CD3epsilon can also be used.  
XX Compositions of the invention have an in trans effect on the replication  
XX of HIV-1, and may be used to treat and prevent HIV-1 infection. The  
XX present sequence represents the chimeric CD4 molecule CD4epsilon1015, which  
XX is specifically claimed for use in compositions of the invention.  
XX  
SQ Sequence 473 AA;  
Query Match 39.9%; Score 1064; DB 7; Length 473;  
Best Local Similarity 100.0%; Pred. No. 9.3e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNRGVPFPHLLLVLTQALLPAATOGKVVLGKKGDTVELTCTASQKKSIOFMKNSNOIK 60  
DB 1 MNRGVPFPHLLLVLTQALLPAATOGKVVLGKKGDTVELTCTASQKKSIOFMKNSNOIK 60  
QY 61 ILNGSGFLTKGSPSKLNRADSRSLMDQGNPPLIIKLIKIEDSPTYICEVEDQKEEYOL 120  
DB 61 ILNGSGFLTKGSPSKLNRADSRSLMDQGNPPLIIKLIKIEDSPTYICEVEDQKEEYOL 120  
QY 121 LVFGLTANSPTHTLQGSGLTLTLESPPGSSPSVOCRSRGRNIOGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSPTHTLQGSGLTLTLESPPGSSPSVOCRSRGRNIOGKTLVSQLELDQSG 180  
QY 181 TWTCTVLONOKKVEFKIDIVVLAFOKAS 208  
DB 181 TWTCTVLONOKKVEFKIDIVVLAFOKAS 208

RESULT 15  
AA59170  
ID AAY59170 standard; protein; 474 AA.  
XX  
XX AAY59170;  
AC AAY59170;  
XX  
XX 14-MAR-2000 (first entry)  
XX  
XX CD4-Ig fusion protein CD4mu.  
XX  
XX HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;  
XX secreted protein; SIV infection; medicament.  
XX  
XX Synthetic.  
XX Homo sapiens.  
XX  
XX CA1340741-C.  
XX  
XX 14-SEP-1999.  
XX  
XX 20-JAN-1989; 89CA-00588749.  
XX  
XX 20-JAN-1989; 89CA-00588749.  
XX  
XX (GEHO ) GEN HOSPITAL CORP.  
XX  
XX Seed B;  
XX  
XX WPI; 2000-061015/06.  
XX  
XX N-PSDB; AA248203.  
XX  
XX New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in  
XX the treatment of HIV or simian immunodeficiency virus infections.  
XX  
XX Example 1; Page 47-53; 89pp; English.  
XX  
XX The invention provides a fusion gene encoding a fusion protein that  
XX comprises an extracellular CD4 DNA sequence or its fragment which binds  
XX to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA  
XX sequence of an Ig heavy or light chain, where the DNA sequence encoding  
XX the variable region has been replaced with the DNA sequence which encodes  
XX extracellular CD4 or its gp120 binding fragment. The fusion protein is  
XX capable of being secreted. The fusion proteins are useful for treating  
XX HIV or SIV infections in animals, preferably humans. They are also useful  
XX for producing medicaments which can be used for treating HIV or SIV  
XX infections in humans. The present sequence represents the fusion protein  
XX CD4mu where the CD4 is linked to human IgG1 at the Met2 site upstream of  
XX the CH1 region  
XX  
SQ Sequence 474 AA;  
Query Match 39.9%; Score 1064; DB 3; Length 474;  
Best Local Similarity 100.0%; Pred. No. 9.3e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNRGVPFPHLLLVLTQALLPAATOGKVVLGKKGDTVELTCTASQKKSIOFMKNSNOIK 60  
DB 1 MNRGVPFPHLLLVLTQALLPAATOGKVVLGKKGDTVELTCTASQKKSIOFMKNSNOIK 60  
QY 61 ILNGSGFLTKGSPSKLNRADSRSLMDQGNPPLIIKLIKIEDSPTYICEVEDQKEEYOL 120  
DB 61 ILNGSGFLTKGSPSKLNRADSRSLMDQGNPPLIIKLIKIEDSPTYICEVEDQKEEYOL 120  
QY 121 LVFGLTANSPTHTLQGSGLTLTLESPPGSSPSVOCRSRGRNIOGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSPTHTLQGSGLTLTLESPPGSSPSVOCRSRGRNIOGKTLVSQLELDQSG 180  
QY 181 TWTCTVLONOKKVEFKIDIVVLAFOKAS 208  
DB 181 TWTCTVLONOKKVEFKIDIVVLAFOKAS 208

Sun Nov 28 09:38:24 2004

us-09-936-702-3.rag

Page 10

Search completed: November 21, 2004, 13:37:17  
Job time : 265.343 secs

---